

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) **EP 0 962 528 A2**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
08.12.1999 Patentblatt 1999/49

(21) Anmeldenummer: **99107852.8**

(22) Anmeldetag: **21.04.1999**

(51) Int. Cl.⁶: **C12N 15/12, C07K 14/705,
C07K 16/28, G01N 33/50,
C12Q 1/68, C12N 5/10,
A01K 67/033, C12N 15/85**

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: **04.05.1998 DE 19819829**

(71) Anmelder:
**Bayer Aktiengesellschaft
51368 Leverkusen (DE)**

(72) Erfinder:
• **Adamczewski, Martin Dr.
51067 Köln (DE)**
• **Oellers, Nadja Dr.
50670 Köln (DE)**
• **Schulte, Thomas Dr.
51061 Köln (DE)**

Bemerkungen:

Der Anmelder hat nachträglich ein Sequenzprotokoll eingereicht und erklärt, dass dieses keine neuen Angaben enthält.

(54) **Nukleisäuren, die für Acetylcholinrezeptor-Untereinheiten von Insekten kodieren**

(57) Die Erfindung betrifft Nukleinsäuren, die für Acetylcholinrezeptor-Untereinheiten von Insekten kodieren, die entsprechenden Polypeptide, sowie Verfahren zum Auffinden neuer Wirkstoffe für den Pflanzenschutz.

EP 0 962 528 A2

Beschreibung

[0001] Die Erfindung betrifft insbesondere Nukleinsäuren, die für Acetylcholinrezeptor-Untereinheiten von Insekten kodieren.

[0002] Nikotinische Acetylcholinrezeptoren sind ligandengesteuerte Ionenkanäle, die eine Rolle bei der Neurotransmission im Tierreich spielen. Die Bindung von Acetylcholin oder anderen Agonisten an den Rezeptor verursacht eine vorübergehende Öffnung des Kanals und gestattet den Durchstrom von Kationen. Man nimmt an, daß ein Rezeptor aus fünf Untereinheiten besteht, die sich um eine Pore gruppieren. Jede dieser Untereinheiten ist ein Protein, das aus einem extrazellulären N-terminalen Teil besteht, gefolgt von drei Transmembranregionen, einem intrazellulären Teil, sowie einer vierten Transmembranregion und einem kurzen extrazellulären C-terminalen Teil (Changeux et al. 1992).

[0003] Acetylcholinrezeptoren sind vor allem in Wirbeltieren gut untersucht. Anhand ihrer anatomischen Lokalisierung und ihrer funktionellen Eigenschaften (Leitungseigenschaften des Kanals, Desensibilisierung, Sensitivität gegenüber Agonisten und Antagonisten, sowie gegenüber Toxinen wie z.B. α -Bungarotoxin) lassen sich hier drei Gruppen unterscheiden. Die Einteilung korreliert mit der molekularen Zusammensetzung der Rezeptoren. Es gibt heterooligomere

Rezeptoren mit der Untereinheitenzusammensetzung $\alpha_2\beta\gamma\delta$, die im Muskel vorkommen (Noda et al. 1982, Claudio et al. 1983, Devillers-Thiery et al. 1983, Noda et al. 1983a, b), heterooligomere Rezeptoren, die Untereinheiten aus der Gruppe $\alpha 2 - \alpha 6$ und $\beta 2 - \beta 4$ enthalten, und die im Nervensystem vorkommen (Wada et al. 1988, Schoepfer et al. 1990, Cockcroft et al. 1991, Heinemann et al. 1997), sowie homooligomere Rezeptoren, die Untereinheiten aus der Gruppe $\alpha 7 - \alpha 9$ enthalten, und die ebenfalls im Nervensystem vorkommen (Lindstrom et al. 1997, Elgoyhen et al. 1997). Diese

Einteilung wird auch durch eine Betrachtung der Verwandtschaft der Gensequenzen der verschiedenen Untereinheiten gestützt. Typischerweise sind die Sequenzen funktionell homologer Untereinheiten verschiedener Spezies ähnlicher als Sequenzen von Untereinheiten aus verschiedenen Gruppen, aber der gleichen Spezies. So weist z.B. die muskuläre α -Untereinheit der Ratte 78 % identische und 84 % ähnliche Aminosäuren auf mit der des elektrischen Rochens *Torpedo californica*, aber nur 48 % Identität und 59 % Ähnlichkeit mit der $\alpha 2$ -Untereinheit (heterooligomer, neuronal) der Ratte, und 36 % Identität und 45 % Ähnlichkeit mit der $\alpha 7$ -Untereinheit (homooligomer, neuronal) der Ratte. Weiterhin sind die Gensequenzen aller bekannten Acetylcholinrezeptor-Untereinheiten nicht nur untereinander in gewissem Maße ähnlich, sondern auch mit denen einiger anderer ligandengesteuerter Ionenkanäle (z.B. den Serotoninrezeptoren vom Typ 5HT₃, den GABA-gesteuerten Chloridkanälen, den Glycin-gesteuerten Chloridkanälen). Man geht daher davon aus, daß alle diese Rezeptoren von einem gemeinsamen Vorläufer abstammen und ordnet sie in eine Supergenfamilie ein (Ortells et al. 1995).

[0004] In Insekten ist Acetylcholin der wichtigste exzitatorische Neurotransmitter des zentralen Nervensystems. Dementsprechend lassen sich Acetylcholinrezeptoren an Präparaten zentraler Ganglien aus Insekten elektrophysiologisch nachweisen. Der Nachweis gelingt sowohl an post- als auch an präsynaptischen Nervenendigungen, sowie an den Zellkörpern von Interneuronen, Motoneuronen und modulatorischen Neuronen (Breer et al. 1987, Buckingham et al. 1997). Unter den Rezeptoren gibt es solche, die durch α -Bungarotoxin inhibiert werden, und solche, die insensitive sind (Schloß et al. 1988). Die Acetylcholinrezeptoren sind außerdem der molekulare Angriffspunkt wichtiger natürlicher (z.B. Nikotin) und synthetischer Insektizide (z.B. Chloronikotinyne).

[0005] Die Gensequenz einer Anzahl von nikotinischen Acetylcholinrezeptoren der Insekten ist bereits bekannt. So sind in *Drosophila melanogaster* die Sequenzen fünf verschiedener Untereinheiten beschrieben (Bossy et al. 1988, Hermanns-Borgmeyer et al. 1986, Sawruk et al. 1990a, 1990b, Schulz et al. unveröffentlicht, EMBL accession number Y15593), in *Locusta migratoria* ebenfalls fünf (Stetzer et al. unveröffentlicht, EMBL accession numbers AJ000390 - AJ000393), in *Schistocerca gregaria* eine (Marshall et al. 1990), in *Myzus persicae* zwei (Sgard et al. unveröffentlicht, EMBL accession number X81887 und X81888), in *Manduca sexta* eine Sequenz (Eastham et al. 1997). Zudem ist eine Reihe von partiellen Gensequenzen aus *Drosophila melanogaster* als sog. expressed sequence tags charakterisiert worden (Genbank accession numbers AA540687, AA698155, AA697710, AA697326). Die hohe Ähnlichkeit einzelner Sequenzen mit solchen aus anderen Insekten legt nahe, daß es sich bei diesen Untereinheiten um funktionelle Homologe handelt.

[0006] Es ist von großer praktischer Bedeutung, beispielsweise für die Suche nach neuen Insektiziden, neue Untereinheiten von Acetylcholinrezeptoren aus Insekten bereitzustellen, wobei besonders solche von Interesse sind, die sich von den bekannten stärker unterscheiden, als dies zwischen funktionellen Homologen der Fall ist.

[0007] Der vorliegenden Erfindung liegt somit insbesondere die Aufgabe zugrunde, Nukleinsäuren zur Verfügung zu stellen, die neue Acetylcholinrezeptor-Untereinheiten von Insekten kodieren.

[0008] Die Aufgabe wird gelöst durch die Bereitstellung von Nukleinsäuren umfassend eine Sequenz ausgewählt aus

(a) den Sequenzen gemäß SEQ ID NO: 1, SEQ ID NO: 3 oder SEQ ID NO: 5,

(b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,

(c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren in 2 x SSC bei 60°C, bevorzugt in 0,5 x SSC bei 60°C, besonders bevorzugt in 0,2 x SSC bei 60°C (Sambrook et al. 1989),

(d) Sequenzen, welche eine zumindest 70 %ige Identität zu den unter (a) definierten Sequenzen zwischen Position 1295 und Position 2195 aus SEQ ID NO: 1 oder zwischen Position 432 und Position 1318 aus SEQ ID NO: 3 oder zwischen Position 154 und Position 1123 aus SEQ ID NO: 5 aufweisen,

(e) Sequenzen, welche zu den unter (a) definierten Sequenzen komplementär sind und

(f) Sequenzen, welche aufgrund der Degeneration des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.

[0009] Der Grad der Identität der Nukleinsäuresequenzen wird vorzugsweise bestimmt mit Hilfe des Programms GAP aus dem Programmpaket GCG, Version 9.1 unter Standardeinstellungen (Devereux et al. 1984).

[0010] Die vorliegende Erfindung begründet sich auf den überraschenden Befund, daß Insekten Gene besitzen, die für Untereinheiten von insbesondere homooligomeren Acetylcholinrezeptoren kodieren.

[0011] Gegenstand der Erfindung sind weiterhin Vektoren, die zumindest eine der erfindungsgemäßen Nukleinsäuren enthalten. Als Vektoren können alle in molekularbiologischen Laboratorien verwendete Plasmide, Phasmide, Cosmide, YACs oder künstliche Chromosomen verwendet werden. Für die Expression der erfindungsgemäßen Nukleinsäuren können diese mit üblichen regulatorischen Sequenzen verknüpft werden. Die Auswahl solcher regulatorischen Sequenzen ist davon abhängig, ob zur Expression pro- oder eukaryotische Zellen bzw. Zellfreie Systeme verwendet werden. Besonders bevorzugt als Expressionskontrollsequenz sind z.B. der frühe oder späte Promotor des SV40 oder des Adenovirus, des Cytomegalovirus, das lac-System, das trp-System, die Haupt-Operator- und Promotorregionen des Phagen lambda, die Kontrollregionen des fd-Hüllproteins, der Promotor der 3-Phosphoglyceratkinase, der Promotor der Sauren Phosphatase und der Promotor des α -Mating-Faktors der Hefe.

[0012] Zur Expression der erfindungsgemäßen Nukleinsäuren können diese in geeignete Wirtszellen eingebracht werden. Als Wirtszellen eignen sich sowohl prokaryotische Zellen, vorzugsweise E.coli, als auch eukaryotische Zellen, vorzugsweise Säuger- oder Insektenzellen. Weitere Beispiele für geeignete einzellige Wirtszellen sind: Pseudomonas, Bacillus, Streptomyces, Hefen, HEK-293, Schneider S2, CHO-, COS1-, COS7-, Zellen, Pflanzenzellen in Zellkultur sowie Amphibienzellen, insbesondere Oocyten.

[0013] Gegenstand der vorliegenden Erfindung sind auch die Polypeptide, die von den erfindungsgemäßen Nukleinsäuren kodiert werden, sowie die daraus aufgebauten Acetylcholinrezeptoren, bevorzugt homooligomere Acetylcholinrezeptoren.

[0014] Zur Herstellung der Polypeptide, die von den erfindungsgemäßen Nukleinsäuren kodiert werden, können Wirtszellen, die zumindest eine der erfindungsgemäßen Nukleinsäuren enthalten, unter geeigneten Bedingungen kultiviert werden. Die gewünschten Polypeptide können danach auf übliche Weise aus den Zellen oder dem Kulturmedium isoliert werden.

[0015] Weiterhin sind Antikörper Gegenstand der Erfindung, die spezifisch an die vorstehend genannten Polypeptide bzw. Rezeptoren binden. Die Herstellung solcher Antikörper erfolgt auf die übliche Weise. Beispielsweise können solche Antikörper produziert werden durch die Injektion eines substantiell immunkompetenten Wirts mit einer für die Antikörperproduktion effektiven Menge eines erfindungsgemäßen Acetylcholinrezeptor-Polypeptids oder eines Fragments davon und durch nachfolgende Gewinnung dieses Antikörpers. Weiterhin läßt sich in an sich bekannter Weise eine immortalisierte Zelllinie erhalten, die monoklonale Antikörper produziert. Die Antikörper können gegebenenfalls mit einem Nachweisreagenz markiert sein. Bevorzugte Beispiele für ein solches Nachweis-Reagenz sind Enzyme, radioaktiv markierte Elemente, fluoreszierende Chemikalien oder Biotin. Anstelle des vollständigen Antikörpers können auch Fragmente eingesetzt werden, die die gewünschten spezifischen Bindungseigenschaften besitzen.

[0016] Die erfindungsgemäßen Nukleinsäuren können insbesondere zur Herstellung transgener Invertebraten verwendet werden. Diese können in Testsysteme eingesetzt werden, die auf einer vom Wildtyp abweichenden Expression der erfindungsgemäßen Rezeptoren oder Varianten hiervon basieren. Ferner fallen hierunter sämtliche transgenen Invertebraten, bei denen durch die Modifikation anderer Gene oder Genkontrollsequenzen (Promotoren) eine Veränderung der Expression der erfindungsgemäßen Rezeptoren oder deren Varianten eintritt.

[0017] Die Herstellung der transgenen Invertebraten erfolgt beispielsweise in Drosophila melanogaster durch P-Element vermittelten Gentransfer (Hay et al., 1997) oder in Caenorhabditis elegans durch Transposon vermittelten Gentransfer (z.B. durch Tc1, Plasterk, 1996).

[0018] Gegenstand der Erfindung sind somit auch transgene Invertebraten, die zumindest eine der erfindungsgemäßen Nukleinsäuren enthalten, vorzugsweise transgene Invertebraten der Arten Drosophila melanogaster oder Caenorhabditis elegans, sowie deren transgene Nachkommen. Vorzugsweise enthalten die transgenen Invertebraten die erfindungsgemäßen Rezeptoren in einer vom Wildtyp abweichenden Form.

[0019] Die erfindungsgemäßen Nukleinsäuren können auf die übliche Weise hergestellt werden. Beispielsweise können die Nukleinsäuremoleküle vollständig chemisch synthetisiert werden. Man kann auch nur kurze Stücke der erfindungsgemäßen Sequenzen chemisch synthetisieren und solche Oligonukleotide radioaktiv oder mit einem Fluoreszenzfarbstoff markieren. Die markierten Oligonukleotide können verwendet werden, um ausgehend von Insekten-mRNA hergestellte cDNA-Banken zu durchsuchen. Klone, an die die markierten Oligonukleotide hybridisieren ("positive Klone"), werden zur Isolierung der betreffenden DNA ausgewählt. Nach der Charakterisierung der isolierten DNA erhält man auf einfache Weise die erfindungsgemäßen Nukleinsäuren.

[0020] Die erfindungsgemäßen Nukleinsäuren können auch mittels PCR-Verfahren unter Verwendung chemisch synthetisierter Oligonukleotide hergestellt werden.

[0021] Die erfindungsgemäßen Nukleinsäuren können zur Isolierung und Charakterisierung der regulatorischen Regionen, die natürlicherweise benachbart zu der kodierenden Region vorkommen, verwendet werden. Solche regulatorischen Regionen sind somit ebenfalls Gegenstand der vorliegenden Erfindung.

[0022] Mit Hilfe der erfindungsgemäßen Nukleinsäuren können neue Wirkstoffe für den Pflanzenschutz identifiziert werden, z.B. Verbindungen, welche als Modulatoren, insbesondere als Agonisten oder Antagonisten, die Leitungseigenschaften der erfindungsgemäßen Acetylcholinrezeptoren verändern. Dazu wird ein rekombinantes DNA-Molekül, das zumindest ein erfindungsgemäßes Nukleinsäuremolekül umfaßt, in eine geeignete Wirtszelle eingebracht. Die Wirtszelle wird in Gegenwart einer Verbindung oder einer Probe, welche eine Vielzahl von Verbindungen umfaßt, unter Bedingungen kultiviert, die die Expression der erfindungsgemäßen Rezeptoren erlauben. Eine Veränderung der Rezeptoreigenschaften kann - wie nachstehend in Beispiel 2 beschrieben - detektiert werden. Auf diese Weise ist es möglich, insektizide Substanzen aufzufinden.

[0023] Die erfindungsgemäßen Nukleinsäuren ermöglichen auch das Auffinden von Verbindungen, die an die erfindungsgemäßen Rezeptoren binden. Diese können ebenfalls als Insektizide auf Pflanzen angewandt werden. Beispielsweise werden Wirtszellen, die die erfindungsgemäßen Nukleinsäuren enthalten und die entsprechenden Rezeptoren bzw. Polypeptide exprimieren oder die Genprodukte selbst mit einer Verbindung oder einem Gemisch von Verbindungen unter Bedingungen in Kontakt gebracht, die die Wechselwirkung zumindest einer Verbindung mit den Wirtszellen, den Rezeptoren oder den einzelnen Polypeptiden erlauben.

[0024] Unter Verwendung von Wirtszellen oder transgenen Invertebraten, die die erfindungsgemäßen Nukleinsäuren enthalten, ist es auch möglich, Substanzen aufzufinden, welche die Expression der Rezeptoren verändern.

[0025] Die vorstehend beschriebenen erfindungsgemäßen Nukleinsäuren, Vektoren und regulatorischen Regionen können außerdem zum Auffinden von Genen verwendet werden, die für Polypeptide kodieren, welche am Aufbau von funktionell ähnlichen Acetylcholinrezeptoren in Insekten beteiligt sind. Unter funktionell ähnlichen Rezeptoren werden gemäß der vorliegenden Erfindung Rezeptoren verstanden, die Polypeptide umfassen, welche sich zwar hinsichtlich der Aminosäuresequenz von den hierin beschriebenen Polypeptiden unterscheiden, aber im wesentlichen dieselben Funktionen haben.

Erläuterungen zum Sequenzprotokoll und zu den Figuren:

[0026]

SEQ ID NO: 1 zeigt die Nukleotidsequenz der isolierten Da7 cDNA, beginnend mit Position 1 und endend mit Position 2886. SEQ ID NO: 1 und SEQ ID NO: 2 zeigen ferner die Aminosäuresequenzen des von der Da7 cDNA Sequenz abgeleiteten Proteins.

SEQ ID NO: 3 zeigt die Nukleotidsequenz der isolierten Hva7-1 cDNA, beginnend mit Position 1 und endend mit Position 3700. SEQ ID NO: 3 und SEQ ID NO: 4 zeigen ferner die Aminosäuresequenzen des von der Hva7-1 cDNA Sequenz abgeleiteten Proteins.

SEQ ID NO: 5 zeigt die Nukleotidsequenz der isolierten Hva7-2 cDNA, beginnend mit Position 1 und endend mit Position 3109. SEQ ID NO: 5 und SEQ ID NO: 6 zeigen ferner die Aminosäuresequenzen des von der Hva7-2 cDNA Sequenz abgeleiteten Proteins.

[0027] Figur 1 zeigt den Anstieg des intrazellulären Calciums in gentechnisch veränderten Zellen gemäß Beispiel 2 nach Gabe von Nikotin. Zellen wurden mit Fura-2-acetoxymethylester (5 - 10 µM in serumfreiem Minimal Essentiellem Medium mit 1 % Rinderserumalbumin and 5 mM Calciumchlorid) beladen, mit N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (5 mM HEPES) gepufferter Tyrode-Lösung gewaschen und unter einem Fluoreszenzmikroskop (Nikon Diaphot) abwechselnd mit Licht der Wellenlänge 340 nm und 380 nm bestrahlt. Ein Meßpunkt entspricht einem Paar von Videobildern bei beiden Wellenlängen (Belichtungszeit pro Bild 100 ms). Der Zeitabstand von zwei Meßpunkten beträgt 3 s. Nach Aufnahme von 8 Bildern (Meßpunkt 4.0) wurde Nikotin auf eine Endkonzentration von 500 µM zuge-

geben, und die Meßreihe fortgesetzt. Die Fluoreszenzintensität der Zellen bei Bestrahlung mit Licht der Wellenlänge 380 nm wurde durch die entsprechende Intensität bei 340 nm geteilt und so das Verhältnis ("Ratio") gebildet.

Beispiele:

Beispiel 1

Isolierung der beschriebenen Polynukleotidsequenzen

[0028] Die Manipulation von Polynukleotiden erfolgte nach Standardmethoden der rekombinanten DNA Technologie (Sambrook, et al., 1989). Die bioinformatische Bearbeitung von Nukleotid- und Proteinsequenzen erfolgten mit dem Programmpaket GCG Version 9.1 (GCG Genetics Computer Group, Inc., Madison Wisconsin, USA).

Partielle Polynukleotidsequenzen

[0029] Aus Proteinsequenzen von Genen, bei denen ihre Fähigkeit homooligomere Acetylcholinrezeptoren auszubilden bekannt war, wurden durch Sequenzvergleiche ("Clustalw") Bereiche identifiziert, aus denen durch Rücktranslation der Codons degenerierte Oligonukleotide abgeleitet wurden. Insgesamt wurden 5 solcher Oligonukleotidpaare für die Polymerasekettenreaktion (PCR) ausgewählt. Nur eine Kombination (vide infra) ergab ein Produkt sowohl aus *Heliothis*-cDNA als auch aus *Drosophila*-cDNA.

[0030] RNA wurde aus gesamten *Heliothis virescens*-Embryonen (kurz vor Schlupf) mittels Trizol-Reagenz (Gibco BRL, nach Angaben des Herstellers) isoliert. In gleicher Weise wurde mit *Drosophila*-Embryonen (24 h bei 25°C) verfahren. 10 µg dieser RNAs wurden in eine cDNA-Erststrangsynthese (Superscript Präamplifizierungssystem für die cDNA-Erststrangsynthese, Gibco BRL, nach Angaben des Herstellers, 45°C Reaktionstemperatur) eingesetzt.

[0031] Anschließend wurden jeweils 1/100 der o.g. Erststrang-cDNA in eine Polymerasekettenreaktion (PCR) mit den Oligonukleotiden alpha7-1s: (5'-GAYGTIGAYGARAARAAYCA-3') und alpha7-2a: (5'-CYYTCRTCIGCRCTRTTTRTA-3') eingesetzt (Taq DNA Polymerase, rekombinant, Gibco BRL). Die PCR-Parameter waren wie folgt: Hva7-1 und Hva7-2: 94°C, 2 min; 35 mal (94°C, 45 s; 50°C, 30 s; 72°C, 60 s) sowie Da7: 96°C, 2 min; 35 mal (96°C, 45 s; 50°C, 30 s; 72°C, 60 s). Hieraus ergab sich jeweils eine im Agarosegel (1 %) erkennbare Bande von ca. 0,2 kb sowohl bei *Drosophila*-cDNA als auch bei *Heliothis*-cDNA. Nach Subklonierung der DNA-Fragmente mittels SrfScript (Stratagene) und Bestimmung der DNA-Sequenz, zeigte sich, daß aus *Heliothis*-cDNA zwei verschiedene DNA-Fragmente amplifiziert worden waren; diese waren 228-11 = Hva7-1 (partiell, mit 165 bp) und 228-8 = Hva7-2 (partiell, mit 171 bp). Aus *Drosophila*-cDNA wurde nur ein DNA-Fragment isoliert; dieses war 248-5 = Da7 (partiell, mit 150 bp).

Isolierung von poly A enthaltender RNA aus *Heliothis virescens*-Gewebe und Konstruktion der cDNA-Bibliotheken

[0032] Die RNA für die cDNA-Bibliothek I wurde aus gesamten *Heliothis virescens*-Embryonen (kurz vor Schlupf) mittels Trizol-Reagenz (Gibco BRL, nach Angaben des Herstellers) isoliert. Die RNA für die cDNA-Bibliothek II wurde aus gesamten Kopfganglien von 500 *Heliothis virescens*-Larven (Stadien 4-5) mittels Trizol-Reagenz (Gibco BRL, nach Angaben des Herstellers) isoliert. Aus diesen RNAs wurden nun die poly A enthaltenden RNAs durch Reinigung über Dyna Beads 280 (Dynal) isoliert. 5 µg dieser poly A enthaltenden RNAs wurden anschließend in die Konstruktion der cDNA-Bibliotheken I und II mit dem λ-ZAPExpress Vektor eingesetzt (cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit und ZAP-cDNA Gigapack III Gold Cloning Kit, alle Stratagene). In Abweichung von den Angaben des Herstellers wurde zur cDNA-Synthese die Reverse Transkriptase Superscript (Gibco BRL) bei einer Synthesetemperatur von 45°C verwendet. Außerdem wurde auf die Zugabe radioaktiv markierter Desoxynukleosidtriphosphate verzichtet. Desweiteren wurden die synthetisierten cDNAs nicht über das im Kit enthaltene Gelfiltrationsmedium, sondern über Size Sep 400 Spun Columns (Pharmacia) fraktioniert.

Vollständige Polynukleotidsequenzen

[0033] Mit Ausnahme der ersten Screening Runde bei der Isolierung des Hva7-1 Klon, erfolgten alle Screens mit Hilfe des DIG Systems (alle Reagenzien und Verbrauchsmaterialien Boehringer Mannheim, nach Angaben im "The DIG System User's Guide for Filter Hybridization", Boehringer Mannheim). Die eingesetzten DNA-Sonden wurden durch PCR mittels Digoxigenin markiertem dUTP präpariert. Die Hybridisierungen erfolgten in DIG Easy Hyb (Boehringer Mannheim) bei 42°C über Nacht. Der Nachweis markierter DNA auf Nylonmembranen geschah durch Chemolumineszenz (CDP-Star, Boehringer Mannheim) unter Verwendung von Röntgenfilmen (Hyperfilm MP, Amersham). Die isolierten Genbankplasmide wurden zur Identifikation mittels T3 und T7 Primer ansequenziert (ABI Prism Dye Terminator Cycle Sequencing Kit, ABI, mit ABI Prism 310 Genetic Analyzer). Die Bestimmung der vollständigen Polynukleo-

EP 0 962 528 A2

tidsequenzen in Hva7-1, Hva7-2 und Da7 erfolgte durch Primer Walking mittels Cycle Sequencing als Auftragssequenzierung bei der Firma Qiagen, Hilden.

a. Isolierung des Da7 Klons

[0034] 10⁶ Phagen einer Drosophila melanogaster-cDNA-Bibliothek in λ Phagen (Canton-S embryo, 2-14 Stunden, in Uni-ZAP XR Vektor, Stratagene) wurden einem Screening mit DIG markiertem 248-5 als Sonde unterzogen (nach Angaben des Herstellers Stratagene). Die maximale Stringenz beim Waschen der Filter betrug: 0,2 x SSC; 0,1 % SDS; 42°C; 2 x 15 min. Es konnte ein Klon isoliert werden (Klon 432-1), dessen Insert eine Größe von 2940 bp aufwies (Da7, SEQ ID NO: 1). Der größte offene Leserahmen dieser Sequenz beginnt bei der Position 372 der dargestellten Sequenz und endet bei Position 1822. Das hieraus abgeleitete Polypeptid umfaßt 770 Aminosäuren (SEQ ID NO: 2) und besitzt ein errechnetes Molekulargewicht von 87,01 kD.

b. Isolierung des Hva7-1 Klons

[0035] In das Screening gingen 10⁶ Phagen der Heliothis virescens-Embryo-cDNA-Bibliothek (Bibliothek I) ein. Die erste von drei Screening Runden fand unter Verwendung α -³²P markierter 228-11 DNA als Sonde statt. Die Hybridisierung der Sonde an die Filter erfolgte in Quickhyb (Stratagene) bei 68°C für eine Stunde. Anschließend wurden die Filter zwei mal je 15 min bei Raumtemperatur in 2 x SSC; 0,1 % SDS und 2 mal je 30 min bei 42°C in 0,1 x SSC; 0,1 % SDS gewaschen. Die Detektion hybridisierter Sonden erfolgte durch Autoradiographie mit XR Röntgenfilmen (Kodak) unter Verwendung von Verstärkerfolien (Amersham) bei -80°C über Nacht. Die zwei weiteren Screening Runden erfolgten unter Verwendung des DIG Systems (Boehringer Mannheim).

[0036] Der in diesem Screen isolierte Klon 241-5 enthielt ein Insert von 3630 bp. Dieses Insert (Hva7-1, SEQ ID NO: 3) besitzt einen längsten offenen Leserahmen, der bei Position 335 der dargestellten Nukleinsäuresequenz beginnt und bei Position 1821 endet. Das hieraus abgeleitete Polypeptid umfaßt 496 Aminosäuren (SEQ ID NO: 4) und besitzt ein errechnetes Molekulargewicht von 56,36 kD.

c. Isolierung des Hva7-2 Klons

[0037] In das Screening gingen 10⁶ Phagen der Heliothis virescens-Ganglien-cDNA-Bibliothek (Bibliothek II) ein. Als Sonde wurde Dig markierte 228-8 DNA verwendet. Die maximale Stringenz beim Waschen der Filter betrug: 0,1 x SSC; 0,1% SDS; 42°C; 2 x 15 min.

[0038] Der in diesem Screen isolierte Klon 241-5 enthielt ein Insert von 3630 bp. Dieses Insert (Hva7-2, SEQ ID NO: 5) besitzt einen längsten offenen Leserahmen, der bei Position 95 der dargestellten Nukleinsäuresequenz beginnt und bei Position 1598 endet. Das hieraus abgeleitete Polypeptid umfaßt 501 Aminosäuren (SEQ ID NO: 6) und besitzt ein errechnetes Molekulargewicht von 56,71 kD.

Beispiel 2

Generierung der Expressionskonstrukte

a. Da7

[0039] Mittels Polymerasekettenreaktion (PCR) wurde der Sequenzbereich von Position 372 bis Position 2681 aus SEQ ID NO: 1 amplifiziert. Hierzu wurden Desoxyoligonukleotide mit den Sequenzen GCGAATTCACCACCATGAAAAATGCACAACTG sowie CGAGACAATAATATGTGGTGCCTCGAG verwendet. Als DNA-Polymerase wurde die Pfu-Polymerase von Stratagene nach Angaben des Herstellers verwendet. Nach erfolgter Amplifikation wurde das generierte Stück mit den Restriktionsendonukleasen Eco RI und Xho I verdaut und in einen ebenfalls Eco RI und Xho I verdauten Vektor pcDNA3.1/Zeo (Invitrogen) einkloniert.

b. Hva7-1

[0040] Mittels Polymerasekettenreaktion (PCR) wurde der Sequenzbereich von Position 335 bis Position 1822 aus SEQ ID NO: 3 amplifiziert. Hierzu wurden Desoxyoligonukleotide mit den Sequenzen GCAAGCTTACCACCATGGGAGGTAGAGCTAGACGCTCGCAC sowie GCCTCGAGCGACACCATGATGTGTGGCGC verwendet. Als DNA-Polymerase wurde die Pfu-Polymerase von Stratagene nach Angaben des Herstellers verwendet. Nach erfolgter Amplifikation wurde das generierte Stück mit den Restriktionsendonukleasen HindIII und Xho I verdaut und in einen ebenfalls HindIII und Xho I verdauten Vektor pcDNA3.1/Zeo (Invitrogen) einkloniert.

c. Hva7-2

[0041] Mittels Polymerasekettenreaktion (PCR) wurde der Sequenzbereich von Position 95 bis Position 1597 aus SEQ ID NO: 5 amplifiziert. Hierzu wurden Desoxyoligonukleotide mit den Sequenzen
 5 GCAAGCGCCGCTATGGCCCCCTATGTTG sowie TTGCACGATGATATGCGGTGCCTCGAGCG verwendet. Als DNA-Polymerase wurde die Pfu-Polymerase von Stratagene nach Angaben des Herstellers verwendet. Nach erfolgter Amplifikation wurde das generierte Stück mit den Restriktionsendonukleasen HindIII und Xho I verdaut und in einen ebenfalls HindIII und XhoI verdauten Vektor pcDNA3.1/Zeo (Invitrogen) einkloniert.

10 d. Hva7-1 / 5HT₃ sowie Hva7-2 / 5HT₃ Chimären

[0042] Durch die Methode der Overlap Extension (Jespersen et al. 1997) wurde jeweils der Bereich von Position 335 bis Position 1036 aus SEQ ID NO: 3 (Hva7-1/5HT₃ Chimäre) sowie der Bereich von Position 95 bis Position 763 aus SEQ ID NO: 5 (Hva7-2/5HT₃ Chimäre) mit dem Bereich von Position 778 bis Position 1521 aus der Mus musculus 5-HT₃ Rezeptor-cDNA (Sequenz in EMBL Datenbank: M774425) fusioniert. Die beiden Fragmente wurden anschließend mittels TA Cloning (Invitrogen, nach Angaben des Herstellers) in den pcDNA3.1/Zeo Vektor kloniert. Konstrukte mit korrekter Orientierung der beiden Fragmente im Vektor wurden durch Sequenzierung mit dem T7 Primer (Invitrogen) identifiziert.

20 Zellkultur und Gentransfer

[0043] HEK293-Zellen, die die α -Untereinheit eines L-Typ Ca-Kanals exprimieren (Zong et al. 1995, Stetzer et al. 1996), wurden in Dulbeccos Modified Eagles Medium und 10 % foetalem Kälberserum bei 5 % CO₂ und 20°C bis 37°C kultiviert. Für den Gentransfer wurde FuGENE 6 (Boehringer Mannheim GmbH, Mannheim, Deutschland) nach Angaben des Herstellers verwendet. 24 h bis 48 h nach dem Gentransfer wurden die Zellen in verschiedenen Dichten in Mikrotiterplatten ausgesät. Gentechnisch veränderte Zellen wurden durch Wachstum in Dulbeccos Modified Eagles Medium und 10 % foetalem Kälberserum und 150 - 500 μ g/ml Zeocin während 3 bis 4 Wochen selektioniert. Resistente Einzelkione wurden wie unten beschrieben analysiert.

30 Fura-2-Messungen

[0044] Die Veränderungen der intrazellulären Calcium-Konzentration wurden mit Fura-2 gemessen. Eine Stammlösung mit 2 mM Fura-2-acetoxymethylester (Sigma) in Dimethylsulfoxid (DMSO) wurde auf eine Endkonzentration von 5 - 10 μ M in serumfreiem Minimal Essentiellm Medium (MEM, Gibco) mit 1% Rinderserumalbumin und 5 mM Calciumchlorid verdünnt. Die Zellen wurden in einer Mikrotiterplatte in dieser Lösung 45 bis 60 min lang inkubiert. Anschließend wurden die Zellen zweimal in N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (5 mM HEPES) gepufferter Tyrode-Lösung (HEPES-gepufferte Salzlösung mit 130 mM NaCl, 5 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂, 5 mM NaHCO₃, 10 mM Glucose) gewaschen. 100 μ l Tyrodepuffer wurde in die Vertiefungen der Mikrotiterplatte gegeben und die Zellen wurden unter einem Fluoreszenzmikroskop (Nikon Diaphot) abwechselnd mit Licht der Wellenlänge 340 nm und 380 nm bestrahlt. Eine Serie von Videobildern (Belichtungszeit pro Bild 100 ms) wurde mit Pausen von 3 Sekunden aufgenommen und als digitalisierte Bilder in einem Bildanalyse-Computer gespeichert (Leica, Quantimet 570). Nach Aufnahme von 8 Bildern (Meßpunkt 4.0 in Fig. 1) wurde Nikotin auf eine Endkonzentration von 500 μ M zugegeben, und die Meßreihe fortgesetzt. Die Fluoreszenzintensität der Zellen bei Bestrahlung mit Licht der Wellenlänge 380 nm wurde durch die entsprechende Intensität bei 340 nm geteilt und so ein Verhältnis gebildet, das den relativen Anstieg der Calcium-Konzentration darstellt (Grynkiewicz et al. 1985).

Literatur:

[0045]

- 50 Bossy et al. (1988) Conservation of neural nicotinic acetylcholine receptors from Drosophila to vertebrate central nervous systems, EMBO J. 7, 611-618
- 55 Breer et al. (1987) Molecular properties and functions of insect acetylcholine receptors, J. Insect Physiol. 33, 771-790
- Buckingham et al. (1997) Imidacloprid actions on insect neuronal acetylcholine receptors, J. Exp. Biol. 200, 2685-2692

EP 0 962 528 A2

Changeux et al. (1992) The functional architecture of the nicotinic acetylcholine receptor explored by affinity labeling and site-directed mutagenesis, *Quarterly Review of Biophysics* 25, 395-432

5 Claudio et al. (1983) Nucleotide and deduced amino acid sequences of *Torpedo californica* acetylcholine receptor γ subunit, *Proc. Natl. Acad. Sci. USA* 80, 1111-1115

Devereux et al. (1984), *Nucleic Acids Research* 12, 387

10 Devillers-Thiery et al. (1983) Complete mRNA coding sequence of the acetylcholine binding α -subunit of *Torpedo marmorata* acetylcholine receptor: a model for the transmembrane organization of the polypeptide chain, *Proc. Natl. Acad. Sci. USA* 80, 2067-2071

Elgoyhen et al. (1997) US Pat. No. 5,683,912

15 Eastham et al. (1998) Characterisation of a nicotinic acetylcholine receptor from the insect *Manduca sexta*, *Eur. J. Neurosci* 10, 879-889

Grynkiewicz et al. (1985) A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, *J Biol Chem.* 260, 3440-3450

20 Hay et al. (1997), P element insertion-dependent gene activation in the *Drosophila* eye, *Proceedings of The National Academy of Sciences of The United States of America* 94 (10), 5195-5200

25 Hermans-Borgmeyer et al. (1986) Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from *Drosophila* EMBO J. 5, 1503-1508

Heinemann et al. (1997) US Pat. No 5,591,590

30 Jespersen et al. (1997) Efficient Non-PCR-Mediated Overlap Extension of PCR Fragments by Exonuclease "End Polishing", *Biotechniques*, 23, 48-52

Lindstrom et al. (1997) US Pat. No. 5,599,709

35 Marshall et al. (1990) Sequence and functional expression of a single α subunit of an insect nicotinic acetylcholine receptor, *EMBO J.* 9, 4391-4398

Noda et al. (1982), Primary structure of α -subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence, *Nature* 299, 793-797

40 Noda et al. (1983a), Primary structures of β - and δ -subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequences, *Nature* 301, 251-255

Noda et al. (1983b), Structural homology of *Torpedo californica* acetylcholine receptor subunits, *Nature* 302, 528-532

45 Ortells et al. (1995), Evolutionary history of the ligand-gated ion-channel superfamily of receptors, *Trends in Neuroscience* 18, 121-127

50 Plasterk (1996), The Tc1/mariner transposon family, *Transposable Elements/Current Topics in Microbiology and Immunology* 204, 125-143

Sambrook et al. (1989), *Molecular Cloning, A Laboratory Manual*, 2nd ed. Cold Spring Harbour Press

55 Sawruk et al. (1990a), EMBO J. 9, 2671-2677 Heterogeneity of *Drosophila* nicotinic acetylcholine receptors: SAD, a novel developmentally regulated α -subunit

Sawruk et al. (1990b), SBD, a novel structural subunit of the *Drosophila* nicotinic acetylcholine receptor, shares its genomic localization with two α -subunits, *FEBS Lett.* 273, 177-181

EP 0 962 528 A2

Schloß et al. (1988), Neuronal acetylcholine receptors of *Drosophila*: the ARD protein is a component of a high-affinity α -bungarotoxin binding complex, EMBO J 7, 2889-2984

5 Stetzer et al. (1996) Stable expression in HEK-293 cells of the rat $\alpha 3/\beta 4$ subtype of neuronal nicotinic acetylcholine receptor, FEBS Lett. 397, 39-44

Zong et al. (1995) On the regulation of the expressed L-type calcium channel by cAMP-dependent phosphorylation, Pflügers Arch. - Eur. J. Physiol. 430, 340-347.

EP 0 962 528 A2

SEQUENZPROTOKOLL

5 <110> Bayer Aktiengesellschaft

<120> Nukleinsäuren, die für
Acetylcholinrezeptor-Untereinheiten von Insekten
kodieren

10 <130> Le A 33 020-EP

<140> 99107852.8

15 <141> 1999-04-21

<150> DE 198 19 829.9

<151> 1998-05-04

20 <160> 6

<170> PatentIn Ver. 2.1

<210> 1

25 <211> 2886

<212> DNA

<213> Drosophila melanogaster

<220>

30 <221> CDS

<222> (372)..(2681)

<400> 1

35 ggcacgagaa aaagttgtgg tataaacttt tattgtagga aaacgcataa aaataataga 60

aaaacgctct tcgggttgta aagaaaataa gaagacaaaa gaaagacatg aaaacggtgc 120

aaacaataaa gcatatactt gccatattga tataaaggga aatcgtgaaa aggcggtgaa 180

40 aatttcgtaa gattagttgg tattaagggc agcccatgca cacagctaaa aagggaacta 240

aaaaaacccc gcacagaaca atgaaagctg cagcagctgg ataaggccga caaaaccgaa 300

45 aattatatta ttgtaatcta gtagagagca gacaacatat ccgctggcaa caaccaacac 360

cgaaagagac t atg aaa aat gca caa ctg aaa ctg act gaa gtt gac gat 410

Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp

1 5 10

50 gat gag ctg tgg ctg gca gta aga tta gcg cac tgc agc agc aac ttt 458

Asp Glu Leu Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe

55

EP 0 962 528 A2

| | 15 | 20 | 25 | |
|----|---|-----|-----|------|
| 5 | agc agc agt agc agc aca aga acc acc agc agc aac cag agg cac aac | | | 506 |
| | Ser Ser Ser Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn | | | |
| | 30 | 35 | 40 | 45 |
| 10 | cag caa ctc aca aca ctg caa cca agg agc tta agt aca aaa cac cac | | | 554 |
| | Gln Gln Leu Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His | | | |
| | 50 | 55 | 60 | |
| 15 | agc aac att gca agc gag cag cac aat agc cag caa cag gag cca gca | | | 602 |
| | Ser Asn Ile Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala | | | |
| | 65 | 70 | 75 | |
| 20 | tcg aag gac gag gat gta gcc aac cac ggt aga agc aat gac cag cag | | | 650 |
| | Ser Lys Asp Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln | | | |
| | 80 | 85 | 90 | |
| 25 | acg cat ctg caa cag cta gac agc agc aac atg ttg tcg cca aag aca | | | 698 |
| | Thr His Leu Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr | | | |
| | 95 | 100 | 105 | |
| 30 | gcc gca gca gca act gct gcc ggc gat gaa gca aca acc caa caa cca | | | 746 |
| | Ala Ala Ala Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro | | | |
| | 110 | 115 | 120 | 125 |
| 35 | aca aac ata aga ctg tgt gca cgc aag cga caa cga ttg cgt cgc cga | | | 794 |
| | Thr Asn Ile Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg | | | |
| | 130 | 135 | 140 | |
| 40 | cga aaa aga aaa cca gca acc cca aac gaa aca gat atc aag aaa caa | | | 842 |
| | Arg Lys Arg Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln | | | |
| | 145 | 150 | 155 | |
| 45 | cag caa ctt agc atg cct ccc ttc aaa acg cgc aaa tcc acg gac acc | | | 890 |
| | Gln Gln Leu Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr | | | |
| | 160 | 165 | 170 | |
| 50 | tac agc aca cca gca gca aca acc agc tgt ccg aca gcc acc tac atg | | | 938 |
| | Tyr Ser Thr Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met | | | |
| | 175 | 180 | 185 | |
| 55 | caa tgt cga gcc agc gac aat gag ttc agt att ccg ata tcg aga cat | | | 986 |
| | Gln Cys Arg Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His | | | |
| | 190 | 195 | 200 | 205 |
| 60 | gat aga gta tcc acg gcc aca ttc gcc tgg gtg ttg cat gtg ctg cag | | | 1034 |
| | Asp Arg Val Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln | | | |

EP 0 962 528 A2

| | 210 | 215 | 220 | |
|----|---|-----|-----|------|
| 5 | gtg ctg ctc gtg tcg ctg caa cag tgg caa ctt cac gtg caa cag cga | | | 1082 |
| | Val Leu Leu Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg | | | |
| | 225 | 230 | 235 | |
| 10 | tcg gtg cta ctg ttc aga agg atc gca gcg agc acc atc gcc ttc att | | | 1130 |
| | Ser Val Leu Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile | | | |
| | 240 | 245 | 250 | |
| 15 | tcc tat tta ggc agc ttt gca gcg caa ctg aaa aat agc agc agc agc | | | 1178 |
| | Ser Tyr Leu Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser | | | |
| | 255 | 260 | 265 | |
| 20 | agt agc agc agc aac agc agc aac aac agc agc acg caa ata tta aac | | | 1226 |
| | Ser Ser Ser Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn | | | |
| | 270 | 275 | 280 | 285 |
| 25 | gga ctt aat aaa cac tca tgg ata ttt tta ttg ata tat ttg aat tta | | | 1274 |
| | Gly Leu Asn Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu | | | |
| | 290 | 295 | 300 | |
| 30 | tct gct aaa gtt tgc cta gca gga tat cat gaa aag aga ctg tta cac | | | 1322 |
| | Ser Ala Lys Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His | | | |
| | 305 | 310 | 315 | |
| 35 | gat ctt ttg gat cct tat aat aca cta gaa cgt ccc gtt ctc aat gaa | | | 1370 |
| | Asp Leu Leu Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu | | | |
| | 320 | 325 | 330 | |
| 40 | tcg gac ccg tta caa tta agc ttt ggt tta act tta atg caa att atc | | | 1418 |
| | Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile | | | |
| | 335 | 340 | 345 | |
| 45 | gat gtg gac gag aaa aat caa ttg cta gtc act aat gtg tgg tta aaa | | | 1466 |
| | Asp Val Asp Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys | | | |
| | 350 | 355 | 360 | 365 |
| 50 | ctg gag tgg aac gac atg aat ctc cgc tgg aac acc tcc gac tat ggc | | | 1514 |
| | Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly | | | |
| | 370 | 375 | 380 | |
| 55 | gga gtt aag gat ctg cga ata ccg ccg cat cgc atc tgg aag ccg gac | | | 1562 |
| | Gly Val Lys Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp | | | |
| | 385 | 390 | 395 | |
| 60 | gtg ctg atg tac aac agt gcg gat gag gga ttt gac ggc acc tac cag | | | 1610 |
| | Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln | | | |

EP 0 962 528 A2

| | 400 | 405 | 410 | |
|----|---|-----|---------|------|
| 5 | acg aac gtg gtg gtg cgg aac aac ggc tcg tgt cta tac gtt ccg ccg Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro 415 | 420 | 425 | 1658 |
| 10 | ggg atc ttc aag tcg acg tgc aag atc gac atc acg tgg ttc ccc ttc Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe 430 | 435 | 440 445 | 1706 |
| 15 | gat gac cag cgg tgc gag atg aag ttc ggc agt tgg acc tac gac gga Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly 450 | 455 | 460 | 1754 |
| 20 | ttc cag ctg gat tta caa tta caa gat gaa act ggc ggt gat atc agc Phe Gln Leu Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser 465 | 470 | 475 | 1802 |
| 25 | agt tac gtg ctc aac ggc gag tgg gaa cta ctg ggt gtg ccc ggc aaa Ser Tyr Val Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys 480 | 485 | 490 | 1850 |
| 30 | cgt aac gag atc tat tac aac tgc tgc ccg gaa ccc tat ata gac atc Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile 495 | 500 | 505 | 1898 |
| 35 | acc ttc gcc atc atc atc cgc cga cga aca ctg tac tat ttc ttc aac Thr Phe Ala Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn 510 | 515 | 520 525 | 1946 |
| 40 | ctg atc ata cct tgt gta ctg att gcc tcc atg gcc ttg ctc gga ttc Leu Ile Ile Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe 530 | 535 | 540 | 1994 |
| 45 | acc ctg ccg cca gat tcg ggt gaa aaa tta tcg ctg ggt gtt acc atc Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile 545 | 550 | 555 | 2042 |
| 50 | ttg ctc tcg ctg acc gtg ttt ctg aat atg gtt gcc gag aca atg ccg Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro 560 | 565 | 570 | 2090 |
| 55 | gct act tcc gat gcg gtg cca ttg tgg ata cgc atc gtg ttt ttg tgc Ala Thr Ser Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys 575 | 580 | 585 | 2138 |
| | tgg ctg cca tgg ata ttg cga atg agt cgc cca gga cga ccg ctg atc Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile | | | 2186 |

EP 0 962 528 A2

| | 590 | 595 | 600 | 605 | |
|----|---|-----|-----|-----|------|
| 5 | cta gag ttc ccg acc acg ccc tgt tgc gac aca tcc tcc gag cgg aag | | | | 2234 |
| | Leu Glu Phe Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys | | | | |
| | 610 | | 615 | 620 | |
| 10 | cac cag ata ctc tcc gac gtt gag ctg aaa gag cgc tgc tgc aaa tgc | | | | 2282 |
| | His Gln Ile Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser | | | | |
| | 625 | | 630 | 635 | |
| 15 | ctg ctg gcc aac gta cta gac atc gat gat gac ttc cgg cac aat tgt | | | | 2330 |
| | Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys | | | | |
| | 640 | | 645 | 650 | |
| 20 | cgc ccc atg acg ccc ggc gga aca ctg cca cac aac ccg gct ttc tat | | | | 2378 |
| | Arg Pro Met Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr | | | | |
| | 655 | | 660 | 665 | |
| 25 | cgc acg gtt tat gga caa ggc gac gat ggc agc att ggg cca att ggc | | | | 2426 |
| | Arg Thr Val Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly | | | | |
| | 670 | | 675 | 680 | 685 |
| 30 | agc acc cga atg ccg gat gcg gtc acc cat cat acg tgc atc aaa tca | | | | 2474 |
| | Ser Thr Arg Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser | | | | |
| | 690 | | 695 | 700 | |
| 35 | tca act gaa tat gaa tta ggt tta atc tta aag gaa att cgc ttt ata | | | | 2522 |
| | Ser Thr Glu Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile | | | | |
| | 705 | | 710 | 715 | |
| 40 | act gat cag cta cgt aaa gat gac gag tgc aat gac att gcc aat gat | | | | 2570 |
| | Thr Asp Gln Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp | | | | |
| | 720 | | 725 | 730 | |
| 45 | tgg aaa ttt gca gct atg gtc gtt gac aga ctg tgc ctt atc ata ttc | | | | 2618 |
| | Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe | | | | |
| | 735 | | 740 | 745 | |
| 50 | aca atg ttc gca ata tta gcc aca ata gct gta cta cta tgc gca cca | | | | 2666 |
| | Thr Met Phe Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro | | | | |
| | 750 | | 755 | 760 | 765 |
| 55 | cat att att gtc tgc tagccatg ggcgaggtgg ttattgttat tggttttatt | | | | 2721 |
| | His Ile Ile Val Ser | | | | |
| | 770 | | | | |
| | ataaaatcaa tttgttaatt attaaattaa taacgaaact ctttaagtaa attaaaacta | | | | 2781 |

EP 0 962 528 A2

aaaagacact aaaaaagcac aaaaaaatag gaaaatacat gataaaaccc atgaactaaa 2841

5 taatacatcc aagaaaaacc aaaacaaaaa aaaaaaaaaa aaaaa 2886

<210> 2

<211> 770

10 <212> PRT

<213> Drosophila melanogaster

<400> 2

15 Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp Asp Glu Leu
1 5 10 15

Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe Ser Ser Ser
20 25 30

20 Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn Gln Gln Leu
35 40 45

25 Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His Ser Asn Ile
50 55 60

Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala Ser Lys Asp
65 70 75 80

30 Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln Thr His Leu
85 90 95

35 Gln Gln Leu Asp Ser Ser Asn Met Leu Ser Pro Lys Thr Ala Ala Ala
100 105 110

Ala Thr Ala Ala Gly Asp Glu Ala Thr Thr Gln Gln Pro Thr Asn Ile
115 120 125

40 Arg Leu Cys Ala Arg Lys Arg Gln Arg Leu Arg Arg Arg Lys Arg
130 135 140

45 Lys Pro Ala Thr Pro Asn Glu Thr Asp Ile Lys Lys Gln Gln Gln Leu
145 150 155 160

Ser Met Pro Pro Phe Lys Thr Arg Lys Ser Thr Asp Thr Tyr Ser Thr
165 170 175

50 Pro Ala Ala Thr Thr Ser Cys Pro Thr Ala Thr Tyr Met Gln Cys Arg
180 185 190

Ala Ser Asp Asn Glu Phe Ser Ile Pro Ile Ser Arg His Asp Arg Val

55

EP 0 962 528 A2

| | | | |
|----|--|-----|-----|
| | 195 | 200 | 205 |
| 5 | Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln Val Leu Leu 210 215 220 | | |
| | Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg Ser Val Leu 225 230 235 240 | | |
| 10 | Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile Ser Tyr Leu 245 250 255 | | |
| 15 | Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser Ser Ser Ser 260 265 270 | | |
| | Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn Gly Leu Asn 275 280 285 | | |
| 20 | Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu Ser Ala Lys 290 295 300 | | |
| | Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His Asp Leu Leu 305 310 315 320 | | |
| 25 | Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu Ser Asp Pro 325 330 335 | | |
| 30 | Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile Asp Val Asp 340 345 350 | | |
| | Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys Leu Glu Trp 355 360 365 | | |
| 35 | Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly Gly Val Lys 370 375 380 | | |
| 40 | Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp Val Leu Met 385 390 395 400 | | |
| | Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln Thr Asn Val 405 410 415 | | |
| 45 | Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe 420 425 430 | | |
| | Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe Asp Asp Gln 435 440 445 | | |
| 50 | Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly Phe Gln Leu | | |

EP 0 962 528 A2

| | 450 | 455 | 460 |
|----|--|-----|---------|
| 5 | Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser Ser Tyr Val 465 | 470 | 475 480 |
| | Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys Arg Asn Glu 485 | 490 | 495 |
| 10 | Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile Thr Phe Ala 500 | 505 | 510 |
| | Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Ile 515 | 520 | 525 |
| 15 | Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe Thr Leu Pro 530 | 535 | 540 |
| 20 | Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile Leu Leu Ser 545 | 550 | 555 560 |
| | Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro Ala Thr Ser 565 | 570 | 575 |
| 25 | Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys Trp Leu Pro 580 | 585 | 590 |
| 30 | Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile Leu Glu Phe 595 | 600 | 605 |
| | Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys His Gln Ile 610 | 615 | 620 |
| 35 | Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser Leu Leu Ala 625 | 630 | 635 640 |
| 40 | Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys Arg Pro Met 645 | 650 | 655 |
| | Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr Arg Thr Val 660 | 665 | 670 |
| 45 | Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly Ser Thr Arg 675 | 680 | 685 |
| 50 | Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser Ser Thr Glu 690 | 695 | 700 |
| | Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile Thr Asp Gln | | |
| 55 | | | |

EP 0 962 528 A2

705 710 715 720
 5 Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp Trp Lys Phe
 725 730 735
 Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe Thr Met Phe
 10 740 745 750
 Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro His Ile Ile
 755 760 765
 15 Val Ser
 770
 20 <210> 3
 <211> 3700
 <212> DNA
 <213> *Heliothis virescens*
 25 <220>
 <221> CDS
 <222> (335) .. (1822)
 30 <400> 3
 ggcacgagcc gctgccccac ggtcgggcgc actccgctga acaacaatgc tcaaaaacac 60
 gccgtgactc cacacacatc ccctcgggcg agtaggcgat gtttgaggat cggacggcac 120
 35 gcgtggccgt cggcgagcgg tcgtgaacaa gttgcataca tatgaaaacc gtaaaaagat 180
 tgaattttaa gccgatcgtg ttcgatagat cctaatagag aagcgggagt gcggcggttg 240
 gtaggcgggg gtcgagtcgc gcggtcgggg gaaatggcgc ggcgcggggc ggcggcggcg 300
 40 gcggcgcgcg gcgcggcggc gtcgcggcgc tgac atg ggc ggg cgg gcg cgc cgc 355
 Met Gly Gly Arg Ala Arg Arg
 1 5
 45 tcg cac ttg gcg gcg ccc gcg ggc ctg ctg ctg ctg ctg tgc ctg ctc 403
 Ser His Leu Ala Ala Pro Ala Gly Leu Leu Leu Leu Cys Leu Leu
 10 15 20
 50 tgg ccg agg ggg gca cgc tgc ggg tac cac gag aag cgg cta ctg cac 451
 Trp Pro Arg Gly Ala Arg Cys Gly Tyr His Glu Lys Arg Leu Leu His
 25 30 35
 55

EP 0 962 528 A2

| | | |
|----|---|------|
| | cac cta ttg gac cac tac aac gta ctg gag agg ccc gtc gtc aac gag | 499 |
| | His Leu Leu Asp His Tyr Asn Val Leu Glu Arg Pro Val Val Asn Glu | |
| 5 | 40 45 50 55 | |
| | agc gac ccg ctg cag ctc tcc ttc ggc ctc acg ctc atg cag atc atc | 547 |
| | Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile | |
| 10 | 60 65 70 | |
| | gac gtg gac gag aag aac cag ctt tta ata aca aac atc tgg cta aaa | 595 |
| | Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Lys | |
| 15 | 75 80 85 | |
| | cta gag tgg aat gat atg aac ttg agg tgg aac act tca gat ttc ggc | 643 |
| | Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Phe Gly | |
| | 90 95 100 | |
| 20 | ggg gtc aaa gat tta aga gtg cca ccc cac aga cta tgg aaa cca gac | 691 |
| | Gly Val Lys Asp Leu Arg Val Pro Pro His Arg Leu Trp Lys Pro Asp | |
| | 105 110 115 | |
| 25 | gtc ctt atg tac aac agc gcg gac gaa ggg ttc gac agc acg tat cca | 739 |
| | Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Ser Thr Tyr Pro | |
| | 120 125 130 135 | |
| 30 | acg aac gtg gtg gtg cgg aac aac ggc tcg tgt ctg tac gtg ccg ccc | 787 |
| | Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro | |
| | 140 145 150 | |
| 35 | ggc atc ttc aag agc acc tgc aag atc gac atc acc tgg ttc ccc ttc | 835 |
| | Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe | |
| | 155 160 165 | |
| 40 | gac gac caa cga tgc gag atg aag ttt ggc agc tgg act tat gat ggt | 883 |
| | Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly | |
| | 170 175 180 | |
| 45 | tat cag ttg gat cta caa cta cag gat gaa ggg ggc gga gat ata agc | 931 |
| | Tyr Gln Leu Asp Leu Gln Leu Gln Asp Glu Gly Gly Gly Asp Ile Ser | |
| | 185 190 195 | |
| 50 | agt ttt gtc acg aat ggc gaa tgg gag tta ata gga gtc ccc ggc aag | 979 |
| | Ser Phe Val Thr Asn Gly Glu Trp Glu Leu Ile Gly Val Pro Gly Lys | |
| | 200 205 210 215 | |
| 55 | cgc aac gag atc tac tac aac tgt tgt ccg gag cca tac atc gac atc | 1027 |
| | Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile | |
| | 220 225 230 | |

EP 0 962 528 A2

| | | |
|----|--|------|
| 5 | acg ttt gcg gtg gtg atc cgg agg aaa acg ctc tac tac ttc ttc aat Thr Phe Ala Val Val Ile Arg Arg Lys Thr Leu Tyr Tyr Phe Phe Asn | 1075 |
| | 235 240 245 | |
| 10 | ctg atc gtg ccc tgc gtg ctc atc gcc tcc atg gct cta ttg ggg ttc Leu Ile Val Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe | 1123 |
| | 250 255 260 | |
| 15 | acc ttg cct cca gac tcc gga gaa aag ttg tct tta ggt gtg acg ata Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile | 1171 |
| | 265 270 275 | |
| 20 | tta ctg tcg ttg acg gtg ttc ctc aac atg gtg gcg gag acg atg cca Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro | 1219 |
| | 280 285 290 295 | |
| 25 | gcg acg tcg gac gcc gtg ccc ttg ctc ggc acc tac ttc aac tgc atc Ala Thr Ser Asp Ala Val Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile | 1267 |
| | 300 305 310 | |
| 30 | atg ttc atg gtg gct tcc tcc gtc gtc tcc acc ata ctg atc ctc aac Met Phe Met Val Ala Ser Ser Val Val Ser Thr Ile Leu Ile Leu Asn | 1315 |
| | 315 320 325 | |
| 35 | tac cac cac cgg cac gca gac act cac gaa atg agt gat tgg att cgt Tyr His His Arg His Ala Asp Thr His Glu Met Ser Asp Trp Ile Arg | 1363 |
| | 330 335 340 | |
| 40 | tgc gtg ttc ctt tat tgg ctg ccg tgg gtg ctg cgc atg tca cgg ccc Cys Val Phe Leu Tyr Trp Leu Pro Trp Val Leu Arg Met Ser Arg Pro | 1411 |
| | 345 350 355 | |
| 45 | ggc tcg gcg acg acg ccg ccg ccg gcg cgc gta cct ccg ccg ccg gac Gly Ser Ala Thr Thr Pro Pro Pro Ala Arg Val Pro Pro Pro Pro Asp | 1459 |
| | 360 365 370 375 | |
| 50 | ctg gag ctg cgc gag cgc tcc tcc aag tcg ctc cta gcg aac gtg ctc Leu Glu Leu Arg Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu | 1507 |
| | 380 385 390 | |
| 55 | gac atc gat gac gac ttc cgc cac ccg caa gcg cag cag ccg caa tgc Asp Ile Asp Asp Asp Phe Arg His Pro Gln Ala Gln Gln Pro Gln Cys | 1555 |
| | 395 400 405 | |
| 60 | tgc cga tac tac agg ggg ggt gag gag aat ggc gcg ggg ttg gcg gcg Cys Arg Tyr Tyr Arg Gly Gly Glu Glu Asn Gly Ala Gly Leu Ala Ala | 1603 |
| | 410 415 420 | |

EP 0 962 528 A2

5 cac agt tgc ttc ggt gtc gac tac gag ctc tcc ctc att ctg aag gag 1651
 His Ser Cys Ph Gly Val Asp Tyr Glu Leu Ser Leu Ile Leu Lys Glu
 425 430 435

10 att aga gtc atc aca gat cag atg cgc aag gac gac gaa gat gcg gac 1699
 Ile Arg Val Ile Thr Asp Gln Met Arg Lys Asp Asp Glu Asp Ala Asp
 440 445 450 455

15 att tcg cgc gac tgg aag ttc gcc gcc atg gtc gtg gac aga ctg tgc 1747
 Ile Ser Arg Asp Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys
 460 465 470

20 ctt att atc ttt acc ctg ttc aca atc atc gcc acg cta gcc gtg ctg 1795
 Leu Ile Ile Phe Thr Leu Phe Thr Ile Ile Ala Thr Leu Ala Val Leu
 475 480 485

25 ctg tcc gcg cca cac atc atg gtg tcg tagcgaccgc cccgcttgcg 1842
 Leu Ser Ala Pro His Ile Met Val Ser
 490 495

30 gatacgcatg cgaaaagtgc tgtgataccg cgaatatctg ttaagttgtg atgagcgaag 1902
 tggcgcgagc ggtgacgcgc cggcgctcgga gttgcccgcg cctgcctcgc cgcccgcgcc 1962
 cccctgtaga cataagttac cgctgactgc caaccctgta cgttcaacaa ataactgccc 2022
 atccgactaa cgtcttttat ccccttgaaa aattcagcga ttgtgtaccc ctttcttcca 2082
 agaatacaat gacaaatggt cgtcacgctc agtggaaatca atcccgact cttcgcccg 2142
 tatttccctt agggatatgc acgagtttga atgagcgggt ccgtatcaga cgttccgtcc 2202
 ccggaacggt cgtccctgc gataaagtgg cagtacgtgc tatacaggca cttaaggccg 2262
 ccacgccacg gcgcccgggt gcgctcgggc cgcgaaccgc cgaccctcac cgctgcaagt 2322
 ggccacccac tagacaagac tgcggcagaa aatatttgca caaaaacgtc ttccttctta 2382
 ccgatgaacg acctgattcg catttaaaat taaactttgt tagaacttct tcgattcttg 2442
 aaatctattg tacagtttag agtttgggcg gtgaacaat ggccctttgt ttccttcttg 2502
 ttcgattcca tgaatcgtgg ttataatccc tagttttatt ttcggatata tttgtgtcag 2562
 tagctagtat agaactttac aaacaatggt gattcaattg gtacagggtg tgatatgcct 2622
 cgttgtgaac gggtcgata ttgttataaa tggtaaaata cccatggcta tagcttaata 2682

55

EP 0 962 528 A2

aatcggttcgt taaaagttgt agttaaaca atattatattt aataaagtca tatctgggtc 2742
 5 ttccggaacg acttttaca ataatataat tacatattaa tatcacgttt gtactttctt 2802
 ccatacagtt acagtaattc gtatgctgaa aataatatta gtttgtaaaa tttcttctt 2862
 10 cgaaaattta ttcaaacaga tgcgaccatc gtttcaaaca ttacatgta atatagaact 2922
 cattttataa gatatacaac attttataag tacaagaagt tgtaacatga accgggtttt 2982
 cgttacatag aggggtataac acaaagggtgc ctacatattg acagatgcga agcacgatca 3042
 15 gttgataagc acaggtacac tatatcctga catccgacag tctgcccgt cgtctgccac 3102
 actcggaac attcgacagt tcagtttact gtcgcccat catcgattgt taagtttgtt 3162
 20 gttctaactc atcgattca tttcattcaa aaacattgta aacctctcaa ggggaaaacg 3222
 tgttgtaaac agtgagagtg cgcgggtaca accgacacgc gaatgtaccc tcgcaaggct 3282
 cctgtaatgt tttctcttc cgagggtgtg ctgagagtaa tcttagacgg tccgatggaa 3342
 25 gttgaggacc ggatatgatt acaagtcaat gtttttaagt catccgttta tttattgtta 3402
 tatcttctta ccattcgta gaggttgtgt gacgaccgg acggtgggcg ccgcaaccg 3462
 30 cacacgagg gttccatctt tgtattagat ggaagtgtg cggcatctct ccgtcggcaa 3522
 tgggacaacc cgttgctccc aacatttgtt caattgttag ggtaactct gaattgcact 3582
 35 ttgtttatta aatataaacg aatgaaaca aaaaaaaaaa aaaaaactcg agagtacttc 3642
 tagagcggcc gcgggcccac cgattttcca cccgggtggg gtaccagtaa gtgtaccc 3700

40 <210> 4
 <211> 496
 <212> PRT
 <213> *Heliothis virescens*

45 <400> 4
 Met Gly Gly Arg Ala Arg Arg Ser His Leu Ala Ala Pro Ala Gly Leu
 1 5 10 15
 50 Leu Leu Leu Leu Cys Leu Leu Trp Pro Arg Gly Ala Arg Cys Gly Tyr
 20 25 30
 His Glu Lys Arg Leu Leu His His Leu Leu Asp His Tyr Asn Val Leu

EP 0 962 528 A2

| | | | |
|----|---|-----|---------|
| | 35 | 40 | 45 |
| 5 | Glu Arg Pro Val Val Asn Glu Ser Asp Pro Leu Gln Leu Ser Phe Gly | | |
| | 50 | 55 | 60 |
| | Leu Thr Leu Met Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu | | |
| | 65 | 70 | 75 80 |
| 10 | Ile Thr Asn Ile Trp Leu Lys Leu Glu Trp Asn Asp Met Asn Leu Arg | | |
| | 85 | 90 | 95 |
| | Trp Asn Thr Ser Asp Phe Gly Gly Val Lys Asp Leu Arg Val Pro Pro | | |
| 15 | 100 | 105 | 110 |
| | His Arg Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu | | |
| | 115 | 120 | 125 |
| 20 | Gly Phe Asp Ser Thr Tyr Pro Thr Asn Val Val Val Arg Asn Asn Gly | | |
| | 130 | 135 | 140 |
| | Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Ile | | |
| 25 | 145 | 150 | 155 160 |
| | Asp Ile Thr Trp Phe Pro Phe Asp Asp Gln Arg Cys Glu Met Lys Phe | | |
| | 165 | 170 | 175 |
| 30 | Gly Ser Trp Thr Tyr Asp Gly Tyr Gln Leu Asp Leu Gln Leu Gln Asp | | |
| | 180 | 185 | 190 |
| | Glu Gly Gly Gly Asp Ile Ser Ser Phe Val Thr Asn Gly Glu Trp Glu | | |
| 35 | 195 | 200 | 205 |
| | Leu Ile Gly Val Pro Gly Lys Arg Asn Glu Ile Tyr Tyr Asn Cys Cys | | |
| | 210 | 215 | 220 |
| 40 | Pro Glu Pro Tyr Ile Asp Ile Thr Phe Ala Val Val Ile Arg Arg Lys | | |
| | 225 | 230 | 235 240 |
| | Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ala | | |
| | 245 | 250 | 255 |
| 45 | Ser Met Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys | | |
| | 260 | 265 | 270 |
| | Leu Ser Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn | | |
| 50 | 275 | 280 | 285 |
| | Met Val Ala Glu Thr Met Pro Ala Thr Ser Asp Ala Val Pro Leu Leu | | |

EP 0 962 528 A2

| | | | |
|----|---|-----|-------------|
| | 290 | 295 | 300 |
| 5 | Gly Thr Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val | | |
| | 305 | 310 | 315 320 |
| | Ser Thr Ile Leu Ile Leu Asn Tyr His His Arg His Ala Asp Thr His | | |
| 10 | | 325 | 330 335 |
| | Glu Met Ser Asp Trp Ile Arg Cys Val Phe Leu Tyr Trp Leu Pro Trp | | |
| | | 340 | 345 350 |
| 15 | Val Leu Arg Met Ser Arg Pro Gly Ser Ala Thr Thr Pro Pro Pro Ala | | |
| | | 355 | 360 365 |
| | Arg Val Pro Pro Pro Pro Asp Leu Glu Leu Arg Glu Arg Ser Ser Lys | | |
| | | 370 | 375 380 |
| 20 | Ser Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Pro | | |
| | | 385 | 390 395 400 |
| | Gln Ala Gln Gln Pro Gln Cys Cys Arg Tyr Tyr Arg Gly Gly Glu Glu | | |
| 25 | | 405 | 410 415 |
| | Asn Gly Ala Gly Leu Ala Ala His Ser Cys Phe Gly Val Asp Tyr Glu | | |
| | | 420 | 425 430 |
| 30 | Leu Ser Leu Ile Leu Lys Glu Ile Arg Val Ile Thr Asp Gln Met Arg | | |
| | | 435 | 440 445 |
| | Lys Asp Asp Glu Asp Ala Asp Ile Ser Arg Asp Trp Lys Phe Ala Ala | | |
| 35 | | 450 | 455 460 |
| | Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe Thr Leu Phe Thr Ile | | |
| | | 465 | 470 475 480 |
| 40 | Ile Ala Thr Leu Ala Val Leu Leu Ser Ala Pro His Ile Met Val Ser | | |
| | | 485 | 490 495 |
| 45 | <210> 5 | | |
| | <211> 3109 | | |
| | <212> DNA | | |
| | <213> Heliothis virescens | | |
| 50 | <220> | | |
| | <221> CDS | | |
| | <222> (95) .. (1597) | | |
| 55 | | | |

EP 0 962 528 A2

<400> 5

ggcaccgagcc ggccgcacgt tgtcccaggc cgcattgagcg cgccggcgtg ctagcgcagc 60

gtgcgcgggt gtggtatgcc cgcgcgtcgc cgct atg gcc cct atg ttg gcg gcc 115
Met Ala Pro Met Leu Ala Ala

1

5

ttg gcg ctg ctg gct ttg ctg ccc gta tcg gag caa ggt cct cac gag 163
Leu Ala Leu Leu Ala Leu Leu Pro Val Ser Glu Gln Gly Pro His Glu
10 15 20

aag aga ctc ctg aac gcg ttg ctg gcg aac tac aac acc ctg gag cga 211
Lys Arg Leu Leu Asn Ala Leu Leu Ala Asn Tyr Asn Thr Leu Glu Arg
25 30 35

ccg gtg gcc aac gag agc gaa ccg cta gag gtc agg ttc ggc ttg acc 259
Pro Val Ala Asn Glu Ser Glu Pro Leu Glu Val Arg Phe Gly Leu Thr
40 45 50 55

ttg cag caa atc att gac gtg gac gag aag aat caa cta ctt ata acc 307
Leu Gln Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr
60 65 70

aat ata tgg ctg tcg ttg gag tgg aat gac tac aac ctg agg tgg aac 355
Asn Ile Trp Leu Ser Leu Glu Trp Asn Asp Tyr Asn Leu Arg Trp Asn
75 80 85

gac agc gag tat ggc ggg gtc aag gac ctc agg atc acg ccc aac aag 403
Asp Ser Glu Tyr Gly Gly Val Lys Asp Leu Arg Ile Thr Pro Asn Lys
90 95 100

ttg tgg aag ccg gac gtc ctt atg tat aat agt gct gac gag ggt ttt 451
Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe
105 110 115

gac ggg acc tac cag acc aac gtg gtg gtc aga agc ggc ggc agt tgc 499
Asp Gly Thr Tyr Gln Thr Asn Val Val Val Arg Ser Gly Gly Ser Cys
120 125 130 135

ctg tac gtg cca cct ggc ata ttc aag agc aca tgc aag atg gac atc 547
Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Met Asp Ile
140 145 150

gcg tgg ttt ccc ttc gac gac caa cac tgt gat atg aag ttc ggt agc 595
Ala Trp Phe Pro Phe Asp Asp Gln His Cys Asp Met Lys Phe Gly Ser
155 160 165

EP 0 962 528 A2

| | | |
|----|--|------|
| 5 | tgg aca tat gac ggc aat cag ttg gat ctg gtg cta aaa gat gag gca Trp Thr Tyr Asp Gly Asn Gln Leu Asp Leu Val Leu Lys Asp Glu Ala | 643 |
| | 170 175 180 | |
| 10 | ggc ggc gat cta tcg gac ttc ata aca aat ggg gag tgg tat cta ata Gly Gly Asp Leu Ser Asp Phe Ile Thr Asn Gly Glu Trp Tyr Leu Ile | 691 |
| | 185 190 195 | |
| 15 | gga atg cca ggc aaa aag aac aca ata aca tac gcg tgc tgc ccc gag Gly Met Pro Gly Lys Lys Asn Thr Ile Thr Tyr Ala Cys Cys Pro Glu | 739 |
| | 200 205 210 215 | |
| 20 | ccc tac gtg gac gtc acc ttc acc atc atg ata aga aga cga acc ttg Pro Tyr Val Asp Val Thr Phe Thr Ile Met Ile Arg Arg Arg Thr Leu | 787 |
| | 220 225 230 | |
| 25 | tac tac ttc ttc aac ctg atc gtc ccg tgc gtg ctg atc tca tcg atg Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ser Ser Met | 835 |
| | 235 240 245 | |
| 30 | gca ctc ctc ggc ttc aca ctg cca cca gac tcc gga gag aaa ctc aca Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Thr | 883 |
| | 250 255 260 | |
| 35 | ctt gga gtc act att ctt cta tcg ctg acg gtg ttc ctc aac ctg gta Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Leu Val | 931 |
| | 265 270 275 | |
| 40 | gcc gag acc ctg cca cag gtc tcc gac gct atc ccc ctg tta ggg acg Ala Glu Thr Leu Pro Gln Val Ser Asp Ala Ile Pro Leu Leu Gly Thr | 979 |
| | 280 285 290 295 | |
| 45 | tac ttc aat tgc atc atg ttc atg gta gcg tcg tct gtg gta ctg act Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val Leu Thr | 1027 |
| | 300 305 310 | |
| 50 | gtg gtg gta ctc aat tac cac cat cga aca gct gat ata cat gaa atg Val Val Val Leu Asn Tyr His His Arg Thr Ala Asp Ile His Glu Met | 1075 |
| | 315 320 325 | |
| 55 | cca cag tgg ata aaa tca gta ttc cta caa tgg ttg cca tgg ata ctg Pro Gln Trp Ile Lys Ser Val Phe Leu Gln Trp Leu Pro Trp Ile Leu | 1123 |
| | 330 335 340 | |
| 60 | cga atg tcg agg cca ggg aag aag atc acc agg aag act ata atg atg Arg Met Ser Arg Pro Gly Lys Lys Ile Thr Arg Lys Thr Ile Met Met | 1171 |
| | 345 350 355 | |

EP 0 962 528 A2

5 aac acg agg atg agg gag ctg gaa ctg aag gag agg tcg tcg aag tcc 1219
Asn Thr Arg Met Arg Glu Leu Glu Leu Lys Glu Arg Ser Ser Lys Ser
360 365 370 375

10 ttg ctg gcg aat gtt cta gat att gat gat gac ttc aga cac ggc cct 1267
Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Gly Pro
380 385 390

15 ccg cct cct aac agt act gcc tcg acc ggg aat ttg gga cct ggg tgc 1315
Pro Pro Pro Asn Ser Thr Ala Ser Thr Gly Asn Leu Gly Pro Gly Cys
395 400 405

20 tca ata ttc cgc acg gat ttc cgt cgg tcg ttc gtc cgt ccg tcc acg 1363
Ser Ile Phe Arg Thr Asp Phe Arg Arg Ser Phe Val Arg Pro Ser Thr
410 415 420

25 atg gaa gac gtg ggc ggc ggg ctg ggt agc cac cat cgc gag ctg cac 1411
Met Glu Asp Val Gly Gly Gly Leu Gly Ser His His Arg Glu Leu His
425 430 435

30 ctc ata ctg aga gag ctg cag ttc atc acg gcc agg atg aag aag gct 1459
Leu Ile Leu Arg Glu Leu Gln Phe Ile Thr Ala Arg Met Lys Lys Ala
440 445 450 455

35 gat gag gaa gcc gag ctg atc agc gac tgg aag ttt gct gcg atg gtt 1507
Asp Glu Glu Ala Glu Leu Ile Ser Asp Trp Lys Phe Ala Ala Met Val
460 465 470

40 gtt gat agg ttt tgc ctg ttc gtg ttc aca ctt ttc aca atc atc gcg 1555
Val Asp Arg Phe Cys Leu Phe Val Phe Thr Leu Phe Thr Ile Ile Ala
475 480 485

45 aca gta gct gtc ctg tta tcg gca ccg cat atc atc gtg caa 1597
Thr Val Ala Val Leu Leu Ser Ala Pro His Ile Ile Val Gln
490 495 500

50 tgaaccaacc actgagccgg caactccggc gcatgaatga gagaaataat tattagatcg 1657

55 ccgatttgta attataattg ataatgtaat taaattaaat acgtggttga aacgcacacg 1717

tctccataac aaagtcttaa gacattaaat tatgataaat ttacatattg tagttaagtc 1777

gagtgttgat ggaaatttta gccggcgcaa ggagtttcgt gaaggtctgt atatattttt 1837

50 tcttattgtt gtatattgta tcgttggtca tgttttcttt caggaagtga gctttgtact 1897

gtttgtttct tcgatggcag gtgcacttca gttcaggctg aaatttccat taacatttat 1957

55

EP 0 962 528 A2

ttaaacaat gtgatgtga ctaggatgtt atacagataa atgttgacgt gtataatttg 2017
 5 ttaaaataaa caatattaat tactattact aaacgatatt ataaacgaag tactaacgag 2077
 gggtacttta atgggaagaa cgctaagctg gcacagagtt gcattaattt gaaaaaagaa 2137
 10 attacggaaa aaagtttatt gaaaattgaa ctttttgaa ggaaagtaac gtttgatcaa 2197
 aaaagtttgt aaaacgaaag ttcggttctg cgccaatact ggaattaaaa ttctcgtaaa 2257
 tattagggaa aagaagggtcc tttaaaacaa aagatttgaa ccggcatcct tttacaagt 2317
 15 aatgagggat cacagatgat gacaaaaaac cttagggat ataagtaatg tacataatgg 2377
 atcaaatatc ggtagagtca agaatagtta acgatttaag attattccat tcgatattaa 2437
 20 aattcgatta gcgattgtcg ctgcgtctac ttgatacat atcgatttga atcgatattg 2497
 tataaattta gatagatcgg acattagtaa tgagtatgga cgttttaatt tttaaaaaag 2557
 aatgtactac gaagattaaa tccaggaatt gttaaacagt tatggaattg ataagaaatc 2617
 25 aacaattaat acggaaccaa aggtagacta ggtgtagcat caggagattg aattaaaaca 2677
 taaattagga ccgacttaaa tggaacttgc gagtgtattg ataacttttt aatttaaaaa 2737
 30 ctcatgtcg attaaatgga gaataacttt tgatctctcg tatcgataaa tgctcactta 2797
 actatcgata gcgtaatatt ataactgtta gtatategat atgggagtaa gtcactagca 2857
 tcagaaatag tcattaatta ggaatcgggt tgtgttaatg ttatgcttag cgaaaatatt 2917
 35 acaatgctgt tgatatcact aaccatcacg taaccatatt gataaatgt aaatacagaa 2977
 tattgcggtg tgtatttgta tataaatttt agaaaaaaaa aaaaaaaaaa aactcgagag 3037
 40 tacttctaga gcggccgcgg gcccatcgat ttccacccg ggtgggttac caggtaagtg 3097
 tacccaattc gc 3109

45 <210> 6
 <211> 501
 <212> PRT
 <213> *Heliothis virescens*

50 <400> 6
 Met Ala Pro Met Leu Ala Ala Leu Ala Leu Leu Ala Leu Pro Val

EP 0 962 528 A2

| | | | | |
|----|---|-----|-----|-----|
| | 1 | 5 | 10 | 15 |
| 5 | Ser Glu Gln Gly Pro His Glu Lys Arg Leu Leu Asn Ala Leu Leu Ala | 20 | 25 | 30 |
| 10 | Asn Tyr Asn Thr Leu Glu Arg Pro Val Ala Asn Glu Ser Glu Pro Leu | 35 | 40 | 45 |
| 15 | Glu Val Arg Phe Gly Leu Thr Leu Gln Gln Ile Ile Asp Val Asp Glu | 50 | 55 | 60 |
| 20 | Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Ser Leu Glu Trp Asn | 65 | 70 | 75 |
| 25 | Asp Tyr Asn Leu Arg Trp Asn Asp Ser Glu Tyr Gly Gly Val Lys Asp | 85 | 90 | 95 |
| 30 | Leu Arg Ile Thr Pro Asn Lys Leu Trp Lys Pro Asp Val Leu Met Tyr | 100 | 105 | 110 |
| 35 | Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln Thr Asn Val Val | 115 | 120 | 125 |
| 40 | Val Arg Ser Gly Gly Ser Cys Leu Tyr Val Pro Pro Gly Ile Phe Lys | 130 | 135 | 140 |
| 45 | Ser Thr Cys Lys Met Asp Ile Ala Trp Phe Pro Phe Asp Asp Gln His | 145 | 150 | 155 |
| 50 | Cys Asp Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly Asn Gln Leu Asp | 165 | 170 | 175 |
| 55 | Leu Val Leu Lys Asp Glu Ala Gly Gly Asp Leu Ser Asp Phe Ile Thr | 180 | 185 | 190 |
| | Asn Gly Glu Trp Tyr Leu Ile Gly Met Pro Gly Lys Lys Asn Thr Ile | 195 | 200 | 205 |
| | Thr Tyr Ala Cys Cys Pro Glu Pro Tyr Val Asp Val Thr Phe Thr Ile | 210 | 215 | 220 |
| | Met Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn Leu Ile Val Pro | 225 | 230 | 235 |
| | Cys Val Leu Ile Ser Ser Met Ala Leu Leu Gly Phe Thr Leu Pro Pro | 245 | 250 | 255 |
| | Asp Ser Gly Glu Lys Leu Thr Leu Gly Val Thr Ile Leu Leu Ser Leu | | | |

EP 0 962 528 A2

| | 260 | 265 | 270 |
|----|---|-------------------------------------|---------|
| 5 | Thr Val Phe Leu Asn Leu Val | Ala Glu Thr Leu Pro Gln Val Ser Asp | |
| | 275 | 280 | 285 |
| 10 | Ala Ile Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile Met Phe Met Val | | |
| | 290 | 295 | 300 |
| | Ala Ser Ser Val Val Leu Thr Val Val Val Leu Asn Tyr His His Arg | | |
| | 305 | 310 | 315 320 |
| 15 | Thr Ala Asp Ile His Glu Met Pro Gln Trp Ile Lys Ser Val Phe Leu | | |
| | 325 | 330 | 335 |
| 20 | Gln Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Lys Lys Ile | | |
| | 340 | 345 | 350 |
| | Thr Arg Lys Thr Ile Met Met Asn Thr Arg Met Arg Glu Leu Glu Leu | | |
| | 355 | 360 | 365 |
| 25 | Lys Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu Asp Ile Asp | | |
| | 370 | 375 | 380 |
| 30 | Asp Asp Phe Arg His Gly Pro Pro Pro Pro Asn Ser Thr Ala Ser Thr | | |
| | 385 | 390 | 395 400 |
| | Gly Asn Leu Gly Pro Gly Cys Ser Ile Phe Arg Thr Asp Phe Arg Arg | | |
| | 405 | 410 | 415 |
| 35 | Ser Phe Val Arg Pro Ser Thr Met Glu Asp Val Gly Gly Gly Leu Gly | | |
| | 420 | 425 | 430 |
| 40 | Ser His His Arg Glu Leu His Leu Ile Leu Arg Glu Leu Gln Phe Ile | | |
| | 435 | 440 | 445 |
| | Thr Ala Arg Met Lys Lys Ala Asp Glu Glu Ala Glu Leu Ile Ser Asp | | |
| | 450 | 455 | 460 |
| 45 | Trp Lys Phe Ala Ala Met Val Val Asp Arg Phe Cys Leu Phe Val Phe | | |
| | 465 | 470 | 475 480 |
| 50 | Thr Leu Phe Thr Ile Ile Ala Thr Val Ala Val Leu Leu Ser Ala Pro | | |
| | 485 | 490 | 495 |
| 55 | His Ile Ile Val Gln | | |
| | 500 | | |

Patentansprüche

1. Nukleinsäure umfassend eine Sequenz ausgewählt aus

- (a) den Sequenzen gemäß SEQ ID NO: 1, SEQ ID NO: 3 oder SEQ ID NO: 5,
- (b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,
- (c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren in 2 x SSC bei 60°C, bevorzugt in 0,5 x SSC bei 60°C, besonders bevorzugt in 0,2 x SSC bei 60°C,
- (d) Sequenzen, welche eine zumindest 70%ige Identität zu den unter (a) definierten Sequenzen zwischen Position 1295 und Position 2195 aus SEQ ID NO: 1 oder zwischen Position 432 und Position 1318 aus SEQ ID NO: 3 oder zwischen Position 154 und Position 1123 aus SEQ ID NO: 5 aufweisen,
- (e) Sequenzen, welche zu den unter (a) definierten Sequenzen komplementär sind und
- (f) Sequenzen, welche aufgrund der Degeneration des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.

2. Vektor umfassend zumindest eine Nukleinsäure gemäß Anspruch 1.

3. Vektor nach Anspruch 2, dadurch gekennzeichnet, daß je besagte Nukleinsäure funktionell mit regulatorischen Sequenzen verknüpft ist, die die Expression der Nukleinsäure in pro- oder eukaryotischen Zellen gewährleisten.

4. Wirtszelle enthaltend eine Nukleinsäure gemäß Anspruch 1 oder einen Vektor gemäß Anspruch 2 oder 3.

5. Wirtszelle nach Anspruch 4, dadurch gekennzeichnet, daß es sich um eine pro- oder eukaryotische Zelle handelt.

6. Wirtszelle nach Anspruch 5, dadurch gekennzeichnet, daß die prokaryotische Zelle E.coli ist.

7. Wirtszelle nach Anspruch 5, dadurch gekennzeichnet, daß die eukaryotische Zelle eine Säuger- oder Insektenzelle ist.

8. Polypeptid, welches von einer Nukleinsäure gemäß Anspruch 1 kodiert wird.

9. Acetylcholinrezeptor umfassend zumindest ein Polypeptid gemäß Anspruch 8.

10. Verfahren zur Herstellung eines Polypeptids gemäß Anspruch 8 umfassend

(a) das Kultivieren einer Wirtszelle gemäß einem der Ansprüche 4 bis 7 unter Bedingungen, die die Expression der Nukleinsäure gemäß Anspruch 1 gewährleisten, und

(b) die Gewinnung des Polypeptids aus der Zelle oder dem Kulturmedium.

11. Antikörper, welcher spezifisch mit dem Polypeptid gemäß Anspruch 8 oder dem Rezeptor gemäß Anspruch 9 reagiert.

12. Transgener Invertebrat enthaltend eine Nukleinsäure gemäß Anspruch 1.

13. Transgener Invertebrat nach Anspruch 12, dadurch gekennzeichnet, daß es sich um Drosophila melanogaster oder Caenorhabditis elegans handelt.

14. Verfahren zur Herstellung eines transgenen Invertebraten gemäß Anspruch 12 oder 13 umfassend das Einbringen einer Nukleinsäure gemäß Anspruch 1 oder eines Vektors gemäß Anspruch 2 oder 3.

15. Transgene Nachkommen eines Invertebraten gemäß Anspruch 12 oder 13.

EP 0 962 528 A2

16. Verfahren zur Herstellung einer Nukleinsäure gemäß Anspruch 1 umfassend die folgenden Schritte:

(a) Vollständige chemische Synthese aufan sich bekannte Weise oder

(b) chemische Synthese von Oligonukleotiden, Markieren der Oligonukleotide, Hybridisieren der Oligonukleotide an DNA einer InsektencDNA-Bank, Selektieren von positiven Klonen und Isolieren der hybridisierenden DNA aus positiven Klonen oder

(c) chemische Synthese von Oligonukleotiden und Amplifizierung der Ziel-DNA mittels PCR.

17. Regulatorische Region, welche natürlicherweise die Transkription einer Nukleinsäure gemäß Anspruch 1 in Insektenzellen kontrolliert und eine spezifische Expression gewährleistet.

18. Verfahren zum Auffinden neuer Wirkstoffe für den Pflanzenschutz, insbesondere von Verbindungen, welche die Leitungseigenschaften von Rezeptoren gemäß Anspruch 9 verändern, umfassend die folgenden Schritte:

(a) Bereitstellen einer Wirtszelle gemäß einem der Ansprüche 4 bis 7,

(b) Kultivieren der Wirtszelle in der Gegenwart einer Verbindung oder einer Probe, welche eine Vielzahl von Verbindungen umfaßt, und

(c) Detektieren veränderter Rezeptoreigenschaften.

19. Verfahren zum Auffinden einer Verbindung, die an Rezeptoren gemäß Anspruch 9 bindet, umfassend die folgenden Schritte:

(a) Inkontaktbringen einer Wirtszelle gemäß einem der Ansprüche 4 bis 7, eines Polypeptids gemäß Anspruch 8 oder eines Rezeptors gemäß Anspruch 9 mit einer Verbindung oder einem Gemisch von Verbindungen unter Bedingungen, die die Interaktion der Verbindung(en) mit der Wirtszelle, dem Polypeptid oder dem Rezeptor erlauben, und

(b) Bestimmen der Verbindung(en), die spezifisch an die Rezeptoren binden.

20. Verfahren zum Auffinden von Verbindungen, die die Expression von Rezeptoren gemäß Anspruch 9 verändern, umfassend die folgenden Schritte:

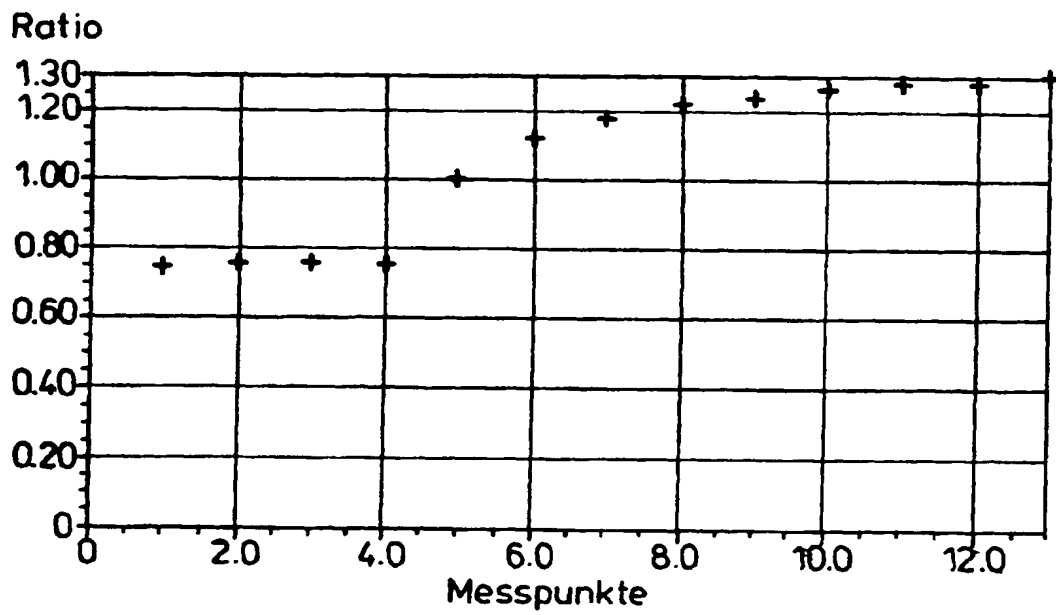
(a) Inkontaktbringen einer Wirtszelle gemäß einem der Ansprüche 4 bis 7 oder eines transgenen Invertebraten gemäß Anspruch 11 oder 12 mit einer Verbindung oder einem Gemisch von Verbindungen,

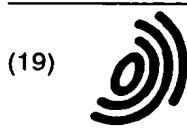
(b) Bestimmen der Rezeptorkonzentration, und

(c) Bestimmen der Verbindung(en), die die Expression des Rezeptors spezifisch beeinflussen.

21. Verwendung zumindest einer Nukleinsäure gemäß Anspruch 1, eines Vektors gemäß Anspruch 2 oder 3, einer regulatorischen Region gemäß Anspruch 16 oder eines Antikörpers gemäß Anspruch 11 zum Auffinden neuer Wirkstoffe für den Pflanzenschutz oder zum Auffinden von Genen, die für Polypeptide kodieren, welche am Aufbau von funktionell ähnlichen Acetylcholinrezeptoren in Insekten beteiligt sind.

Fig. 1





Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) **EP 0 962 528 A3**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(88) Veröffentlichungstag A3:
04.10.2001 Patentblatt 2001/40

(43) Veröffentlichungstag A2:
08.12.1999 Patentblatt 1999/49

(21) Anmeldenummer: 99107852.8

(22) Anmeldetag: 21.04.1999

(51) Int Cl.7: **C12N 15/12**, C07K 14/705,
C07K 16/28, G01N 33/50,
C12Q 1/68, C12N 5/10,
A01K 67/033, C12N 15/85

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 04.05.1998 DE 19819829

(71) Anmelder: **Bayer Aktiengesellschaft**
51368 Leverkusen (DE)

(72) Erfinder:
• **Adamczewski, Martin Dr.**
51067 Köln (DE)
• **Oellers, Nadja Dr.**
50670 Köln (DE)
• **Schulte, Thomas Dr.**
51061 Köln (DE)

(54) **Nukleisäuren, die für Acetylcholinrezeptor-Untereinheiten von Insekten kodieren**

(57) Die Erfindung betrifft Nukleinsäuren, die für Acetylcholinrezeptor-Untereinheiten von Insekten ko-

dieren, die entsprechenden Polypeptide, sowie Verfahren zum Auffinden neuer Wirkstoffe für den Pflanzenschutz.

EP 0 962 528 A3



Europäisches
Patentamt

EUROPÄISCHER TEILRECHERCHENBERICHT

der nach Regel 45 des Europäischen Patent-
Übereinkommens für das weitere Verfahren als
europäischer Recherchenbericht gilt

Nummer der Anmeldung

EP 99 10 7852

| EINSCHLÄGIGE DOKUMENTE | | | |
|---|--|-----------------------------|--|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int.Cl.4) |
| X | WO 96 41876 A (SIBIA NEUROSCIENCES INC ;ELLIOTT KATHRYN J (US); HARPOLD MICHAEL M) 27. Dezember 1996 (1996-12-27) * 100% Identität in einem 15 bp langen überlappenden Bereich zwischen SEQ ID NO 3 aus WO9641876 und SEQ ID NO 3 * * 100% Identität in einem 14 bp langen überlappenden Bereich zwischen SEQ ID NO 11 aus WO9641876 und SEQ ID NO 5 * * 61.1% Identität in einem 635 bp langen überlappenden Bereich zwischen SEQ ID NO 11 aus WO9641876 und SEQ ID NO 3 * | 1-10,16 | C12N15/12 C07K14/705 C07K16/28 G01N33/50 C12Q1/68 C12N5/10 A01K67/033 C12N15/85 |
| X | DATABASE EMBL 'Online! accession: AJ000393, 8. Januar 1998 (1998-01-08) STETZER E: "Locusta migratoria mRNA for nAChR beta subunit" XP002150465 * 58% Identität in einem 1002 bp langen überlappenden Bereich mit SEQ ID NO 5 * * 59.2% Identität in einem 813 bp langen überlappenden Bereich mit SEQ ID NO 1 * | 1-10,16 | |
| | | | RECHERCHIERTE SACHGEBIETE (Int.Cl.8) |
| | | | C12N C07K G01N C12Q A01K |
| UNVOLLSTÄNDIGE RECHERCHE | | | |
| <p>Die Recherchenabteilung ist der Auffassung, daß ein oder mehrere Ansprüche, den Vorschriften des EPU in einem solchen Umfang nicht entspricht bzw. entsprechen, daß eine/volle Ermittlungen über den Stand der Technik für diese Ansprüche nicht, bzw. nur teilweise, möglich sind.</p> <p>Vollständig recherchierte Patentansprüche:</p> <p>Unvollständig recherchierte Patentansprüche:</p> <p>Nicht recherchierte Patentansprüche:</p> <p>Grund für die Beschränkung der Recherche:</p> <p>Siehe Ergänzungsblatt C</p> | | | |
| Recherchenort | | Abschlußdatum der Recherche | |
| DEN HAAG | | 20. Oktober 2000 | |
| | | Prüfer | |
| | | Devijver, K | |
| KATEGORIE DER GENANNTEN DOKUMENTEN | | | |
| <p>X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichttechnische Offenbarung P : Zwischenliteratur</p> <p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentedokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, Übereinstimmendes Dokument</p> | | | |



Europäisches
Patentamt

**UNVOLLSTÄNDIGE RECHERCHE
ERGÄNZUNGSBLATT C**

Nummer der Anmeldung
EP 99 10 7852

Vollständig recherchierte Ansprüche:
1-16, 18-20, 21 (teilweise)

Unvollständig recherchierte Ansprüche:
21

Nicht recherchierte Ansprüche:
17

Grund für die Beschränkung der Recherche:

Die geltenden Patentansprüche 17 und 21, insofern letzterer sich auf Anspruch 17 (und nicht Anspruch 16) rückbezieht, beziehen sich auf eine Nukleinsäure-Region, jeweils charakterisiert durch eine erstrebenswerte Eigenheit oder Eigenschaft, nämlich dass besagte Region eine regulatorische Eigenschaft hat. Die Patentansprüche umfassen daher alle regulatorische Regionen, die diese Eigenheit oder Eigenschaft aufweisen. Dahingegen fehlt in der Patentanmeldung jegliche Stütze durch die Beschreibung im Sinne von Art. 83 EPÜ. Im vorliegenden Fall fehlen den Patentansprüchen die entsprechende Stütze bzw. der Patentanmeldung die nötige Offenbarung in einem solchen Maße, daß eine sinnvolle Recherche für obengenannte Patentansprüche unmöglich erscheint. Desungeachtet fehlt den Patentansprüchen auch die in Art. 84 EPÜ geforderte Klarheit, nachdem in ihnen versucht wird, besagte Nukleinsäure-Region über das jeweils erstrebte Ergebnis zu definieren. Auch dieser Mangel an Klarheit ist dergestalt, dass er eine sinnvolle Recherche unmöglich macht. Daher kann keine sinnvolle Recherche für besagte Patentansprüche durchgeführt werden.

EP 0 962 528 A3



Europäisches
Patentamt

**EUROPÄISCHER
TEILRECHERCHENBERICHT**

Nummer der Anmeldung
EP 99 10 7852

| EINSCHLÄGIGE DOKUMENTE | | | KLASSIFIKATION DER ANMELDUNG (Int.Cl.6) |
|------------------------|--|--------------------|---|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich der maßgeblichen Teile | Betrifft Anspruch | |
| X | DATABASE EMBL 'Online! accession: AC004326, 13. März 1998 (1998-03-13) CELNIKER S E ET AL: "Drosophila melanogaster DNA sequence (P1 DS05899 (D22)), complete sequence" XP002150466 * 98.5% Identität in einem 851 bp langen überlappenden Bereich mit SEQ ID NO 1 * | 1-10, 16 | |
| A | US 5 693 492 A (CULLY DORIS F ET AL) 2. Dezember 1997 (1997-12-02) * das ganze Dokument * | 1-11, 16, 18-21 | |
| | | | RECHERCHIERTE SACHGEBIETE (Int.Cl.6) |
| | | | |
| | | | |

**ANHANG ZUM EUROPÄISCHEN RECHERCHENBERICHT
ÜBER DIE EUROPÄISCHE PATENTANMELDUNG NR.**

EP 99 10 7852

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten europäischen Recherchenbericht angeführten Patentedokumente angegeben.

Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am
Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

20-10-2000

| Im Recherchenbericht angeführtes Patentedokument | Datum der Veröffentlichung | Mitglied(er) der Patentfamilie | Datum der Veröffentlichung |
|---|-------------------------------|-----------------------------------|-------------------------------|
| WO 9641876 A | 27-12-1996 | AU 704927 B | 06-05-1999 |
| | | AU 6266496 A | 09-01-1997 |
| | | CA 2221311 A | 27-12-1996 |
| | | EP 0832230 A | 01-04-1998 |
| | | JP 11509726 T | 31-08-1999 |
| US 5693492 A | 02-12-1997 | CA 2220116 A | 07-11-1996 |
| | | EP 0871703 A | 21-10-1998 |
| | | JP 11506906 T | 22-06-1999 |
| | | WO 9634940 A | 07-11-1996 |

EPO FORM P0481

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82

Nucleic acids which encode insect acetylcholine receptor subunits

The invention relates, in particular, to nucleic acids which encode insect acetylcholine receptor subunits.

Nicotinic acetylcholine receptors are ligand-regulated ion channels which are of importance in neurotransmission in the animal kingdom. The binding of acetylcholine or other agonists to the receptor induces a transient opening of the channel and allows cations to flow through. It is assumed that a receptor consists of five subunits which are grouped around a pore. Each of these subunits is a protein which consists of an extracellular N-terminal moiety followed by three transmembrane regions, an intracellular moiety, a fourth transmembrane region and a short extracellular C-terminal moiety (Changeux et al. 1992).

Acetylcholine receptors are especially well investigated in vertebrates. In this context, three groups can be distinguished on the basis of their anatomical location and their functional properties (conducting properties of the channel, desensitization, and sensitivity towards agonists and antagonists and also towards toxins such as α -bungarotoxin). The classification correlates with the molecular composition of the receptors. There are heterooligomeric receptors having the subunit composition $\alpha_2\beta\gamma\delta$, which are found in muscle (Noda et al. 1982, Claudio et al. 1983, Devillers-Thiery et al. 1983, Noda et al. 1983a, b), heterooligomeric receptors which contain subunits from the α_2 - α_6 and β_2 - β_4 groups and which are found in the nervous system (Wada et al. 1988, Schoepfer et al. 1990, Cockcroft et al. 1991, Heinemann et al. 1997), and also homooligomeric receptors which contain subunits from the α_7 - α_9 group and which are likewise found in the nervous system (Lindstrom et al. 1997, Elgoyhen et al. 1997). This classification is also supported by an examination of the relatedness of the gene sequences of the different subunits. Typically, the sequences of functionally homologous subunits from different species are more similar to each other than are sequences of subunits which are from different groups but from the

same species. Thus, the rat muscle α subunit, for example, exhibits 78% amino acid identity and 84% amino acid similarity with that of the electric ray *Torpedo californica* but only 48% identity and 59% similarity with the rat $\alpha 2$ subunit (heterooligomeric, neuronal) and 36% identity and 45% similarity with the rat $\alpha 7$ subunit (homooligomeric, neuronal). Furthermore, the gene sequences of all the known acetylcholine receptor subunits are to a certain extent similar not only to each other but also to those of some other ligand-regulated ion channels (e.g. the serotonin receptors of the 5HT₃ type, the GABA-regulated chloride channels and the glycine-regulated chloride channels). It is therefore assumed that all these receptors are descended from one common precursor and they are classified into one supergene family (Ortells et al. 1995).

In insects, acetylcholine is the most important excitatory neurotransmitter of the central nervous system. Accordingly, acetylcholine receptors can be detected electrophysiologically in preparations of insect central nervous system ganglia. The receptors are detected both in postsynaptic and presynaptic nerve endings and in the cell bodies of interneurons, motor neurons and modulatory neurons (Breer et al. 1987, Buckingham et al. 1997). Some of the receptors are inhibited by α -bungarotoxin while others are insensitive (Schloß et al. 1988). In addition, the acetylcholine receptors are the molecular point of attack for important natural (e.g. nicotine) and synthetic insecticides (e.g. chloronicotinylns).

The gene sequences of a number of insect nicotinic acetylcholine receptors are already known. Thus, the sequences of five different subunits have been described in *Drosophila melanogaster* (Bossy et al. 1988, Hermanns-Borgmeyer et al. 1986, Sawruk et al. 1990a, 1990b, Schulz et al. Unpublished, EMBL accession number Y15593), while five have likewise been described in *Locusta migratoria* (Stetzer et al. unpublished, EMBL accession numbers AJ000390 - AJ000393), one has been described in *Schistocerca gregaria* (Marshall et al. 1990), two have been described in *Myzus persicae* (Sgard et al. unpublished, EMBL accession number X81887 and X81888), and one has been described in *Manduca sexta* (Eastham et al. 1997). Fur-

thermore, a number of partial gene sequences from *Drosophila melanogaster* have been characterized as so-called expressed sequence tags (Genbank accession numbers AA540687, AA698155, AA697710, AA697326). The fact that individual sequences are very similar to those from other insects suggests that these subunits are functional homologues.

It is of great practical importance to make available new insect acetylcholine receptor subunits, for example for the purpose of searching for novel insecticides, with those subunits which differ from the known subunits to a greater extent than is the case between functional homologues being particularly of interest.

The present invention is consequently based, in particular, on the object of making available nucleic acids which encode novel insect acetylcholine receptor subunits.

This object is achieved by the provision of nucleic acids which comprise a sequence selected from

- (a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,
- (b) part sequences of the sequences defined in (a) which are least 14 base pairs in length,
- (c) sequences which hybridize to the sequences defined in (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C (Sambrook et al. 1989),
- (d) sequences which exhibit at least 70% identity with the sequences defined in (a), between position 1295 and position 2195 in the case of SEQ ID NO: 1, or between position 432 and position 1318 in the case of SEQ ID NO: 3, or between position 154 and position 1123 in the case of SEQ ID NO: 5,

- (e) sequences which are complementary to the sequences defined in (a), and
- (f) sequences which, because of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined in (a) to (d).

5

The degree of identity of the nucleic acid sequences is preferably determined using the GAP program from the GCG program package, Version 9.1 with standard settings (Devereux et al. 1984).

10

The present invention is based on the surprising finding that insects possess genes which encode subunits of, in particular, homooligomeric acetylcholine receptors.

15

The invention furthermore relates to vectors which contain at least one of the novel nucleic acids. All the plasmids, phasmids, cosmids, YACs or artificial chromosomes which are used in molecular biological laboratories can be used as vectors. These vectors can be linked to the usual regulatory sequences for the purpose of expressing the novel nucleic acids. The choice of such regulatory sequences depends on whether prokaryotic or eukaryotic cells, or cell-free systems, are used for the expression. The SV40, adenovirus or cytomegalovirus early or late promoter, the lac system, the trp system, the main operator and promoter regions of phage lambda, the control regions of the fd coat protein, the 3-phosphoglycerate kinase promoter, the acid phosphatase promoter and the yeast α -mating factor promoter are examples of expression control sequences which are particularly preferred.

20

25

In order to be expressed, the nucleic acids according to the invention can be introduced into suitable host cells. Both prokaryotic cells, preferably E.coli, and eukaryotic cells, preferably mammalian or insect cells, are suitable for use as host cells. Other examples of suitable unicellular host cells are: Pseudomonas, Bacillus, Streptomyces, yeasts, HEK-293, Schneider S2, CHO, COS1 and COS7 cells, plant cells in cell culture and also amphibian cells, in particular oocytes.

30

The present invention also relates to polypeptides which are encoded by the nucleic acids according to the invention and also the acetylcholine receptors, preferably homooligomeric acetylcholine receptors, which are synthesized from them.

5 In order to prepare the polypeptides which are encoded by the nucleic acids according to the invention, host cells which contain at least one of the nucleic acids according to the invention can be cultured under suitable conditions. After that, the desired polypeptides can be isolated from the cells or the culture medium in a customary manner.

10 The invention furthermore relates to antibodies which bind specifically to the above-mentioned polypeptides or receptors. These antibodies are prepared in the customary manner. For example, such antibodies can be produced by injecting a substantially immunocompetent host with a quantity of an acetylcholine receptor polypeptide, or a fragment thereof, according to the invention which is effective for producing anti-
15 bodies, and subsequently isolating these antibodies. Furthermore, an immortalized cell line which produces monoclonal antibodies can be obtained in a manner known per se. Where appropriate, the antibodies can be labelled with a detection reagent. Preferred examples of such a detection reagent are enzymes, radioactively labelled elements, fluorescent chemicals or biotin. Instead of the complete antibody, use can also be made
20 of fragments which possess the desired specific binding properties.

The nucleic acids according to the invention can be used, in particular, for producing transgenic invertebrates. These latter can be employed in test systems which are based on an expression of the receptors according to the invention, or variants thereof, which
25 differs from that of the wild type. In addition, this includes all transgenic invertebrates in which a change in the expression of the receptors according to the invention, or their variants, occurs as the result of modifying other genes or gene control sequences (promoters).

The transgenic invertebrates are produced, for example, in *Drosophila melanogaster* by means of P element-mediated gene transfer (Hay et al., 1997) or in *Caenorhabditis elegans* by means of transposon-mediated gene transfer (e.g. using Tc1, Plasterk, 1996).

5 The invention also consequently relates to transgenic invertebrates which contain at least one of the nucleic acid sequences according to the invention, preferably to transgenic invertebrates of the species *Drosophila melanogaster* or *Caenorhabditis elegans*, and to their transgenic progeny. Preferably, the transgenic invertebrates contain the receptors according to the invention in a form which differs from that of the wild type.

10

The nucleic acids according to the invention can be prepared in the customary manner. For example, the nucleic acid molecules can be synthesized entirely chemically. In addition, only short segments of the sequences according to the invention can be synthesized chemically and these oligonucleotides can be labelled radioactively or with a fluorescent dye. The labelled oligonucleotides can be used to screen cDNA libraries prepared from insect mRNA. Clones which hybridize to the labelled oligonucleotides ("positive clones") are selected for isolating the relevant DNA. After the isolated DNA has been characterized, the nucleic acids according to the invention are readily obtained.

15

20

The nucleic acids according to the invention can also be prepared by means of PCR methods using chemically synthesized oligonucleotides.

25 The nucleic acids according to the invention can be used for isolating and characterizing the regulatory regions which occur naturally adjacent to the coding region. Consequently, the present invention also relates to these regulatory regions.

30 The nucleic acids according to the invention can be used to identify novel active compounds for plant protection, such as compounds which, as modulators, in particular as agonists or antagonists, alter the conducting properties of the acetylcholine receptors according to the invention. For this, a recombinant DNA molecule, which encompasses

at least one nucleic acid according to the invention, is introduced into a suitable host cell. The host cell is cultured, in the presence of a compound or a sample which comprises a multiplicity of compounds, under conditions which permit expression of the receptors according to the invention. A change in the receptor properties can be detected, as described below in Example 2. Using this approach, it is possible to discover insecticidal substances.

The nucleic acids according to the invention also make it possible to discover compounds which bind to the receptors according to the invention. These compounds can likewise be used as insecticides on plants. For example, host cells which contain the nucleic acid sequences according to the invention and express the corresponding receptors or polypeptides, or the gene products themselves, are brought into contact with a compound or a mixture of compounds under conditions which permit the interaction of at least one compound with the host cells, receptors or the individual polypeptides.

Host cells or transgenic invertebrates which contain the nucleic acids according to the invention can also be used to discover substances which alter the expression of the receptors.

The above-described nucleic acids, vectors and regulatory regions according to the invention can additionally be used for discovering genes which encode polypeptides which are involved in the synthesis, in insects, of functionally similar acetylcholine receptors. According to the present invention, functionally similar receptors are understood as being receptors which encompass polypeptides which, while differing in their amino acid sequences from the polypeptides described in this present publication, essentially possess the same functions.

Comments on the sequence listing and the figures:

5 SEQ ID NO: 1 shows the nucleotide sequence of the isolated Da7 cDNA, beginning with position 1 and ending with position 2886. SEQ ID NO: 1 and SEQ ID NO: 2 also show the amino acid sequences of the protein deduced from the Da7 cDNA sequence.

10 SEQ ID NO: 3 shows the nucleotide sequence of the isolated Hva7-1 cDNA, beginning with position 1 and ending with position 3700. SEQ ID NO: 3 and SEQ ID NO: 4 also show the amino acid sequences of the protein deduced from the Hva7-1 cDNA sequence.

15 SEQ ID NO: 5 shows the nucleotide sequence of the isolated Hva7-2 cDNA, beginning with position 1 and ending with position 3109. SEQ ID NO: 5 and SEQ ID NO: 6 also show the amino acid sequences of the protein deduced from the Hva7-2 cDNA sequence.

20 Figure 1 shows the increase in intracellular calcium which occurs in cells which have been recombinantly modified as described in Example 2 following the addition of nicotine. Cells were loaded with Fura-2-acetoxymethyl ester (5 - 10 μ M in serum-free minimal essential medium containing 1% bovine serum albumin and 5 mM calcium chloride), washed with Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) and alternately
25 illuminated, under a fluorescence microscope (Nikon Diaphot) with light of 340 nm and 380 nm wavelength. A measurement point corresponds to a pair of video images at the two wavelengths (exposure time per image, 100 ms). The time interval between two measurement points is 3 s. After 8 images had been taken (measurement point 4.0), nicotine was added to a final concentration of 500 μ M and the measurement series was continued. The fluorescence intensity of the cells when illuminated
30

with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm, thereby giving the ratio.

Examples:

5

Example 1

Isolating the described polynucleotide sequences

10 Polynucleotides were manipulated using standard methods of recombinant DNA technology (Sambrook, et al., 1989). The bioinformatic processing of nucleotide and protein sequences was carried out using the GCG program package Version 9.1 (GCG Genetics Computer Group, Inc., Madison Wisconsin, USA).

15 Partial polynucleotide sequences

Sequence comparisons ("Clustalw") were used to identify regions, from which degenerate oligonucleotides were deduced by backtranslating the codons, of protein sequences from genes whose ability to form homooligomeric acetylcholine receptors was known. In all, 5 such oligonucleotide pairs were selected for the polymerase chain reaction (PCR). Only one combination (see below) gave a product both from Heliothis cDNA and from Drosophila cDNA.

25 RNA was isolated from whole Heliothis virescens embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The same procedure was adopted with Drosophila embryos (24 h at 25°C). 10 µg of these RNAs were employed in a first cDNA strand synthesis (Superscript Pre-amplification System for first cDNA strand synthesis, Gibco BRL, in accordance with the manufacturer's instructions, reaction temperature 45°C).

30

Subsequently, 1/100 of the abovementioned first-strand cDNA was in each case employed in a polymerase chain reaction (PCR) using the oligonucleotides alpha7-1s: (5'-GAYGTIGAYGARAARAAYCA-3') and alpha7-2a: (5'-CYYTCRTCIGCRCTRTRTA-3') (recombinant Taq DNA polymerase, Gibco BRL). The PCR parameters were as follows: Hva7-1 and Hva7-2: 94°C, 2 min; 35 times (94°C, 45 s; 50°C, 30 s; 72°C, 60 s) and also Da7: 96°C, 2 min; 35 times (96°C, 45 s; 50°C, 30 s; 72°C, 60 s). In each case, this resulted in a dectable band of approx. 0.2 kb in an agarose gel (1%), both in the case of *Drosophila* cDNA and in the case of *Heliothis* cDNA. After the DNA fragments had been subcloned by means of SrfScript (Stratagene), and their sequences had been determined, it turned out that two different DNA fragments had been amplified from *Heliothis* cDNA; these were 228-11 = Hva7-1 (partial, containing 165 bp) and 228-8 = Hva7-2 (partial, containing 171 bp). Only one DNA fragment was isolated from *Drosophila* cDNA; this was 248-5 = Da7 (partial, containing 150 bp).

15

Isolating poly A-containing RNA from *Heliothis virescens* tissue and constructing the cDNA libraries

The RNA for cDNA library I was isolated from whole *Heliothis virescens* embryos (shortly before hatching) using Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The RNA for cDNA library II was isolated from whole head ganglia from 500 *Heliothis virescens* larvae (stages 4-5) usings Trizol reagent (Gibco BRL, in accordance with the manufacturer's instructions). The poly A-containing RNAs were then isolated from these RNAs by purifying with Dyna Beads 280 (Dynal). 5 µg of these poly A-containing RNAs were subsequently employed in constructing cDNA libraries I and II using the λ-ZAPExpress vector (cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit, all from Stratagene). In a departure from the manufacturer's instructions, Superscript Reverse Transcriptase (Gibco BRL) was used for synthesizing the cDNA at a synthesis temperature of 45°C. In addition, radioactively labelled deoxynucleoside triphosphates were not added. Furthermore, the synthetisized cDNAs were not frac-

30

tionated through the gel filtration medium contained in the kit but instead through Size Sep 400 Spun Columns (Pharmacia).

Complete polynucleotide sequences

5

Apart from the first screening round when isolating the Hva7-1 clone, all the screens were carried out using the DIG system (all reagents and consumables from Boehringer Mannheim, in accordance with the instructions in "The DIG System User's Guide for Filter Hybridization", Boehringer Mannheim). The DNA probes employed were prepared by means of PCR using digoxigenin-labelled dUTP. The hybridizations were carried out at 42°C overnight in DIG Easy Hyb (Boehringer Mannheim). Labelled DNA was detected on nylon membranes by means of chemiluminescence (CDP-Star, Boehringer Mannheim) using X-ray films (Hyperfilm MP, Amersham). Initial partial sequencing of the isolated gene library plasmids was carried out, for identification purposes, using T3 and T7 primers (ABI Prism Dye Terminator Cycle Sequencing Kit, ABI, using an ABI Prism 310 Genetic Analyzer). The complete polynucleotide sequences in Hva7-1, Hva7-2 and Da7 were determined, as a commissioned sequencing carried out by Qiagen, Hilden, by means of primer walking using cycle sequencing.

20

a. Isolating the Da7 clone

10⁶ phages from a *Drosophila melanogaster* cDNA library in λ phages (Canton-S embryo, 2-14 hours, in Uni-ZAP XR vector, Stratagene) were screened using DIG-labelled 248-5 as the probe (in accordance with the manufacturer's (Stratagene) instructions). The maximum stringency when washing the filters was: 0.2 x SSC; 0.1% SDS; 42°C; 2 x 15 min. One clone (clone 432-1) was isolated whose insert had a size of 2940 bp (Da7, SEQ ID NO: 1). The largest open reading frame of this sequence begins at position 372 of the depicted sequence and ends at position 1822. The 770 amino acids polypeptide which is deduced from this (SEQ ID NO: 2) has a calculated molecular weight of 87.01 kD.

30

b. Isolating the Hva7-1 clone

10⁶ phages from the *Heliothis virescens* embryo cDNA library (library I) were included in the screening. The first of three screening rounds took place using α -³²P-labelled 228-11 DNA as the probe. The probe was hybridized to the filters in Quick-hyb (Stratagene) at 68°C for one hour. The filters were then washed twice, for 15 min on each occasion, at room temperature in 2 x SSC; 0.1% SDS and twice, for 30 min on each occasion, at 42°C in 0.1xSSC; 0.1% SDS. Hybridized probes were detected by means of autoradiography, at -80°C overnight, using XR X-ray films (Kodak) and employing intensifying screens (Amersham). The two further screening rounds were carried out using the DIG System (Boehringer Mannheim).

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-1, SEQ ID NO: 3) possesses a longest open reading frame which begins at position 335 of the depicted nucleic acid sequence and ends at position 1821. The 496 amino acids polypeptide which is deduced from this (SEQ ID NO: 4) has a calculated molecular weight of 56.36 kD.

c. Isolating the Hva7-2 clone

10⁶ phages from the *Heliothis virescens* ganglia cDNA library (library II) were included in the screening. Dig-labelled 228-8 DNA was used as the probe. The maximum stringency when washing the filters was: 0.1 x SSC; 0.1% SDS; 42°C; 2 x 15 min.

The clone 241-5, which was isolated in this screen, contained an insert of 3630 bp. This insert (Hva7-2, SEQ ID NO: 5) possesses a longest open reading frame which begins at position 95 of the depicted nucleic acid sequence and ends at position 1598. The 501 amino acids polypeptide which is deduced from this (SEQ ID NO: 6) has a calculated molecular weight of 56.71 kD.

Example 2

Generating the expression constructs

5

a. Da7

10 The sequence region from position 372 to position 2681 of SEQ ID NO: 1 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCGAATTCACCACCATGAAAAATGCACAACTG and CGAGACAATAATATGTGGTGCCTCGAG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following the amplification, the segment which had been generated was digested with the restriction endonucleases Eco RI and Xho I and
15 cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with Eco RI and Xho I.

b. Hva7-1

20 The sequence region from position 335 to position 1822 from SEQ ID NO: 3 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences
GCAAGCTTACCACCATGGGAGGTAGAGCTAGACGCTCGCAC and
GCCTCGAGCGACACCATGATGTGTGGCGC were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the
25 manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

30

Example 2

Generating the expression constructs

5

a. Da7

10

15

The sequence region from position 372 to position 2681 of SEQ ID NO: 1 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCGAATTCACCACCATGAAAAATGCACAACCTG and CGAGACAATAATATGTGGTGCCTCGAG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following the amplification, the segment which had been generated was digested with the restriction endonucleases Eco RI and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with Eco RI and Xho I.

b. Hva7-1

20

25

The sequence region from position 335 to position 1822 from SEQ ID NO: 3 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides having the sequences GCAAGCTTACCACCATGGGAGGTAGAGCTAGACGCTCGCAC and GCCTCGAGCGACACCATGATGTGTGGCGC were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII and Xho I.

30

c. Hva7-2

The sequence region from position 95 to position 1597 from SEQ ID NO: 5 was amplified by means of a polymerase chain reaction (PCR). Deoxyoligonucleotides
5 having the sequences GCAAGCGCCGCTATGGCCCCCTATGTTG and TTGCACGATGATATGCGGTGCCTCGAGCG were used for this. The Pfu polymerase from Stratagene was used as the DNA polymerase in accordance with the manufacturer's instructions. Following amplification, the generated segment was digested with the restriction endonucleases HindIII and Xho I and cloned into a vector, i.e. pcDNA3.1/Zeo (Invitrogen), which had likewise been digested with HindIII
10 and Xho I.

d.Hva7-1 / 5HT₃ and Hva7-2 / 5HT₃ chimaeras

15 The region from position 335 to position 1036 from SEQ ID NO: 3 (Hva7-1/5HT₃ chimaera) and the region from position 95 to position 763 from SEQ ID NO: 5 (Hva7-2/5HT₃ chimaera) was in each case fused to the region from position 778 to position 1521 from the Mus musculus 5-HT₃ receptor cDNA (sequence in EMBL database: M774425) using the method of overlap extension (Jespersen et al. 1997).
20 The two fragments were subsequently cloned into the pcDNA3.1/Zeo vector by means of TA cloning (Invitrogen, in accordance with the manufacturer's instructions). Constructs containing the correct orientation of the two fragments in the vector were identified by sequencing using the T7 primer (Invitrogen).

25 Cell culture and gene transfer

HEK293 cells, which express the α subunit of an L-type Ca channel (Zong et al. 1995, Stetzer et al. 1996), were cultured in Dulbecco's modified Eagle's medium and 10% foetal calf serum at 5% CO₂ and from 20°C to 37°C. FuGENE 6 (Boehringer
30 Mannheim GmbH, Mannheim, Germany) was used for the gene transfer in accordance with the manufacturer's instructions. At from 24 h to 48 h after the gene trans-

fer, the cells were sown at various densities in microtitre plates. Recombinantly altered cells were selected by growth in Dulbecco's modified Eagle's medium and 10% foetal calf serum and 150 - 500 $\mu\text{g/ml}$ of Zeocin/ml over a period of from 3 to 4 weeks. Individual resistant clones were analyzed as described below.

5

Fura-2 measurements

The alterations in the intracellular calcium concentration were measured using Fura-2. A stock solution containing 2 mM Fura-2-acetoxy methyl ester (Sigma) in dimethyl sulphoxide (DMSO) was diluted to a final concentration of 5 - 10 μM in serum-free minimal essential medium (MEM, Gibco) containing 1% bovine serum albumin and 5 mM calcium chloride. The cells were incubated for from 45 to 60 min in this solution in a microtitre plate. The cells were then washed twice in Tyrode solution buffered with N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid) (5 mM HEPES) (HEPES-buffered salt solution containing 130 mM NaCl, 5 mM KCl, 2 mM CaCl_2 , 1 mM MgCl_2 , 5 mM NaHCO_3 , 10 mM glucose). 100 μl Tyrode buffer were added to the wells of the microtitre plate and the cells were illuminated alternately, under a fluorescence microscope (Nikon Diaphot), with light of 340 nm and 380 nm wavelength. A series of video images (exposure time per image 100 ms) were taken with pauses of 3 seconds and stored, as digitalized images, in an image analysis computer (Leica, Quantimet 570). After 8 images had been taken (measurement point 4.0 in Fig. 1), nicotine was added to a final concentration of 500 μM and the measurement series was continued. The fluorescence intensity of the cells when illuminating with light of 380 nm wavelength was divided by the corresponding intensity at 340 nm and in this way a ratio was formed which represents the relative increase in calcium concentration (Grynkiewicz et al. 1985).

10

15

20

25

References:

5 Bossy et al. (1988) Conservation of neural nicotinic acetylcholine receptors from *Drosophila* to vertebrate central nervous systems, *EMBO J.* 7, 611-618

Breer et al. (1987) Molecular properties and functions of insect acetylcholine receptors, *J. Insect Physiol.* 33, 771-790

10 Buckingham et al. (1997) Imidacloprid actions on insect neuronal acetylcholine receptors, *J. Exp. Biol.* 200, 2685-2692

15 Changeux et al. (1992) The functional architecture of the nicotinic acetylcholine receptor explored by affinity labelling and site-directed mutagenesis, *Quarterly Review of Biophysics* 25, 395-432

Claudio et al. (1983) Nucleotide and deduced amino acid sequences of *Torpedo californica* acetylcholine receptor γ subunit, *Proc. Natl. Acad. Sci. USA* 80, 1111-1115

20 Devereux et al. (1984), *Nucleic Acids Research* 12, 387

25 Devillers-Thiery et al. (1983) Complete mRNA coding sequence of the acetylcholine binding α -subunit of *Torpedo marmorata* acetylcholine receptor: a model for the transmembrane organization of the polypeptide chain, *Proc. Natl. Acad. Sci. USA* 80, 2067-2071

Elgoyhen et al. (1997) US Pat. No. 5,683,912

30 Eastham et al. (1998) Characterisation of a nicotinic acetylcholine receptor from the insect *Manduca sexta*, *Eur. J. Neurosci* 10, 879-889

- Grynkiewicz et al. (1985) A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, *J. Biol. Chem.* 260, 3440-3450
- 5 Hay et al. (1997), P element insertion-dependent gene activation in the *Drosophila* eye, *Proceedings of The National Academy of Sciences of The United States of America* 94 (10), 5195-5200
- Hermans-Borgmeyer et al. (1986) Primary structure of a developmentally regulated nicotinic acetylcholine receptor protein from *Drosophila* *EMBO J.* 5, 1503-1508
- 10 Heinemann et al. (1997) US Pat. No 5,591,590
- Jespersen et al. (1997) Efficient Non-PCR-Mediated Overlap Extension of PCR Fragments by Exonuclease "End Polishing", *Biotechniques*, 23, 48-52
- 15 Lindstrom et al. (1997) US Pat. No. 5,599,709
- Marshall et al. (1990) Sequence and functional expression of a single α subunit of an insect nicotinic acetylcholine receptor, *EMBO J.* 9, 4391-4398
- 20 Noda et al. (1982), Primary structure of α -subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence, *Nature* 299, 793-797
- Noda et al. (1983a), Primary structures of β - and δ -subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequences, *Nature* 301, 251-255
- 25 Noda et al. (1983b), Structural homology of *Torpedo californica* acetylcholine receptor subunits, *Nature* 302, 528-532
- 30 Ortells et al. (1995), Evolutionary history of the ligand-gated ion-channel superfamily of receptors, *Trends in Neuroscience* 18, 121-127

- Plasterk (1996), The Tc1/mariner transposon family, *Transposable Elements, Current Topics in Microbiology and Immunology* 204, 125-143
- 5 Sambrook et al. (1989), *Molecular Cloning, A Laboratory Manual*, 2nd ed. Cold Spring Harbour Press
- Sawruk et al. (1990a), *EMBO J.* 9, 2671-2677 Heterogeneity of *Drosophila* nicotinic acetylcholine receptors: SAD, a novel developmentally regulated α -subunit
- 10 Sawruk et al. (1990b), SBD, a novel structural subunit of the *Drosophila* nicotinic acetylcholine receptor, shares its genomic localization with two α -subunits, *FEBS Lett.* 273, 177-181
- 15 Schloß et al. (1988), Neuronal acetylcholine receptors of *Drosophila*: the ARD protein is a component of a high-affinity α -bungarotoxin binding complex, *EMBO J.* 7, 2889-2984
- 20 Stetzer et al. (1996) Stable expression in HEK-293 cells of the rat $\alpha 3/\beta 4$ subtype of neuronal nicotinic acetylcholine receptor, *FEBS Lett.* 397, 39-44
- ~ Zong et al. (1995) On the regulation of the expressed L-type calcium channel by cAMP-dependent phosphorylation, *Pflügers Arch. - Eur. J. Physiol.* 430, 340-347.

Patent Claims

1. Nucleic acid which comprises a sequence selected from

5 (a) the sequences according to SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 5,

(b) parts of sequences which are at least 14 base pairs in length, of the sequences defined under (a),

10 (c) sequences which hybridize with the sequences defined under (a) in 2 x SSC at 60°C, preferably in 0.5 x SSC at 60°C, particularly preferably in 0.2 x SSC at 60°C,

15 (d) sequences which exhibit at least 70% identity with the sequences defined under (a), between position 1295 and position 2195 from SEQ ID NO: 1, or between position 432 and position 1318 from SEQ ID NO: 3, or between position 154 and position 1123 from SEQ ID NO: 5,

20 (e) sequences which are complementary to the sequences defined under (a), and

25 (f) sequences which, on account of the degeneracy of the genetic code, encode the same amino acid sequences as the sequences defined under (a) to (d).

3. Vector according to Claim 2, characterized in that the nucleic acid is functionally linked to regulatory sequences which ensure the expression of the nucleic acid in prokaryotic or eukaryotic cells.
- 5 4. Host cell which contains a nucleic acid according to Claim 1 or a vector according to Claim 2 or 3.
5. Host cell according to Claim 4, characterized in that it is a prokaryotic or eukaryotic cell.
- 10 6. Host cell according to Claim 5, characterized in that the prokaryotic cell is E.coli.
- 15 7. Host cell according to Claim 5, characterized in that the eukaryotic cell is a mammalian cell or an insect cell.
8. Polypeptide which is encoded by a nucleic acid according to Claim 1.
9. Acetylcholine receptor which comprises at least one polypeptide according to Claim 8.
10. Process for preparing a polypeptide according to Claim 8, which comprises
 - 25 (a) culturing a host cell according to one of Claims 4 to 7 under conditions which ensure the expression of the nucleic acid according to Claim 1, and
 - (b) isolating the polypeptide from the cell or the culture medium.
- 30 11. Antibody which reacts specifically with the polypeptide according to Claim 8 or the receptor according to Claim 9.

12. Transgenic invertebrate which contains a nucleic acid according to Claim 1.

13. Transgenic invertebrate according to Claim 12, characterized in that it is
5 *Drosophila melanogaster* or *Caenorhabditis elegans*.

Process for producing a transgenic invertebrate according to Claim 12 or 13,
which comprises introducing a nucleic acid according to Claim 1 or a vector
according to Claim 2 or 3.

10 15. Transgenic progeny of an invertebrate according to Claim 12 or 13.

16. Process for preparing a nucleic acid according to Claim 1, which comprises
the following steps:

15 (a) carrying out an entirely chemical synthesis in a manner known per se,
or

20 (b) chemically synthesizing oligonucleotides, labelling the oligonucleo-
tides, hybridizing the oligonucleotides to the DNA of an insect cDNA
library, selecting positive clones and isolating the hybridizing DNA
from positive clones, or

25 (c) chemically synthesizing oligonucleotides and amplifying the target
DNA by means of PCR.

17. Regulatory region which naturally controls transcription of a nucleic acid
according to Claim 1 in insect cells and ensures specific expression.

18. Process for discovering novel active compounds for plant protection, in particular compounds which alter the conducting properties of receptors according to Claim 9, which comprises the following steps:

- 5 (a) providing a host cell according to one of Claims 4 to 7;
- (b) culturing the host cell in the presence of a compound or a sample which comprises a multiplicity of compounds; and
- 10 (c) detecting altered receptor properties.

19. Process for discovering a compound which binds to receptors according to Claim 9, which encompasses the following steps:

- 15 (a) bringing a host cell according to one of Claims 4 to 7, a polypeptide according to Claim 8, or a receptor according to Claim 9 into contact with a compound or a mixture of compounds under conditions which permit interaction of the compound(s) with the host cell, the polypeptide or the receptor, and
- 20 (b) determining the compound(s) which bind(s) specifically to the receptors.

20. Process for discovering compounds which alter the expression of receptors according to Claim 9, which comprises the following steps:

- 25 (a) bringing a host cell according to one of Claims 4 to 7 or a transgenic invertebrate according to Claim 11 or 12 into contact with a compound or a mixture of compounds,
- 30 (b) determining the receptor concentration, and

(c) determining the compound(s) which specifically influence(s) the expression of the receptor.

- 5 21. Use of at least one nucleic acid according to Claim 1, one vector according to Claim 2 or 3, one regulatory region according to Claim 16 or one antibody according to Claim 11 for discovering novel active compounds for plant protection or for discovering genes which encode polypeptides which are involved in synthesizing functionally similar acetylcholine receptors in insects.
- 10

SEQUENCE LISTING

<110> Bayer Aktiengesellschaft

<120> Nucleic acids which encode
insect acetylcholine receptor subunits

<130> Le A 33 020-Foreign Countries

<140>

<141>

<150> DE-198 19 829.9

<151> 1998-05-04

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 2886

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (372) .. (2681)

<400> 1

ggcacgagaa aaagttgtgg tataaacttt tattgtagga aaacgcataa aaataataga 60

aaaacgctct tcgggttgta aagaaaataa gaagacaaaa gaaagacatg aaaacgttgc 120

aaacaataaa gcatatactt gccatattga tataaaggga aatcgtgaaa aggcggtgaa 180

aatttcgtaa gattagttgg tattaagggc agcccatgca cacagctaaa aagggaacta 240

aaaaaacccc gcacagaaca atgaaagctg cagcagctgg ataaggccga caaaaccgaa 300

aattatatta ttgtaatcta gtagagagca gacaacatat ccgctggcaa caaccaacac 360

cgaaagagac t atg aaa aat gca caa ctg aaa ctg act gaa gtt gac gat 410

Met Lys Asn Ala Gln Leu Lys Leu Thr Glu Val Asp Asp

1

5

10

gat gag ctg tgg ctg gca gta aga tta gcg cac tgc agc agc aac ttt 458

Asp Glu Leu Trp Leu Ala Val Arg Leu Ala His Cys Ser Ser Asn Phe

15

20

25

agc agc agt agc agc aca aga acc acc agc agc aac cag agg cac aac 506

Ser Ser Ser Ser Ser Thr Arg Thr Thr Ser Ser Asn Gln Arg His Asn

30

35

40

45

cag caa ctc aca aca ctg caa cca agg agc tta agt aca aaa cac cac 554

Gln Gln Leu Thr Thr Leu Gln Pro Arg Ser Leu Ser Thr Lys His His

50

55

60

agc aac att gca agc gag cag cac aat agc cag caa cag gag cca gca 602

Ser Asn Ile Ala Ser Glu Gln His Asn Ser Gln Gln Gln Glu Pro Ala

65

70

75

tcg aag gac gag gat gta gcc aac cac ggt aga agc aat gac cag cag 650

Ser Lys Asp Glu Asp Val Ala Asn His Gly Arg Ser Asn Asp Gln Gln

Asp Arg Val Ser Thr Ala Thr Phe Ala Trp Val Leu His Val Leu Gln

210

215

220

gtg ctg ctc gtg tcg ctg caa cag tgg caa ctt cac gtg caa cag cga 1082

Val Leu Leu Val Ser Leu Gln Gln Trp Gln Leu His Val Gln Gln Arg

225

230

235

tcg gtg cta ctg ttc aga agg atc gca gcg agc acc atc gcc ttc att 1130

Ser Val Leu Leu Phe Arg Arg Ile Ala Ala Ser Thr Ile Ala Phe Ile

240

245

250

tcc tat tta ggc agc ttt gca gcg caa ctg aaa aat agc agc agc agc 1178

Ser Tyr Leu Gly Ser Phe Ala Ala Gln Leu Lys Asn Ser Ser Ser Ser

255

260

265

agt agc agc agc aac agc agc aac aac agc agc acg caa ata tta aac 1226

Ser Ser Ser Ser Asn Ser Ser Asn Asn Ser Ser Thr Gln Ile Leu Asn

270

275

280

285

gga ctt aat aaa cac tca tgg ata ttt tta ttg ata tat ttg aat tta 1274

Gly Leu Asn Lys His Ser Trp Ile Phe Leu Leu Ile Tyr Leu Asn Leu

290

295

300

tct gct aaa gtt tgc cta gca gga tat cat gaa aag aga ctg tta cac 1322

Ser Ala Lys Val Cys Leu Ala Gly Tyr His Glu Lys Arg Leu Leu His

305

310

315

gat ctt ttg gat cct tat aat aca cta gaa cgt ccc gtt ctc aat gaa 1370

Asp Leu Leu Asp Pro Tyr Asn Thr Leu Glu Arg Pro Val Leu Asn Glu

320

325

330

| | |
|---|------|
| tcg gac ccg tta caa tta agc ttt ggt tta act tta atg caa att atc | 1418 |
| Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile | |
| 335 340 345 | |
| gat gtg gac gag aaa aat caa ttg cta gtc act aat gtg tgg tta aaa | 1466 |
| Asp Val Asp Glu Lys Asn Gln Leu Leu Val Thr Asn Val Trp Leu Lys | |
| 350 355 360 365 | |
| ctg gag tgg aac gac atg aat ctc cgc tgg aac acc tcc gac tat ggc | 1514 |
| Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Tyr Gly | |
| 370 375 380 | |
| gga gtt aag gat ctg cga ata ccg ccg cat cgc atc tgg aag ccg gac | 1562 |
| Gly Val Lys Asp Leu Arg Ile Pro Pro His Arg Ile Trp Lys Pro Asp | |
| 385 390 395 | |
| gtg ctg atg tac aac agt gcg gat gag gga ttt gac ggc acc tac cag | 1610 |
| Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Gly Thr Tyr Gln | |
| 400 405 410 | |
| acg aac gtg gtg gtg cgg aac aac ggc tcg tgt cta tac gtt ccg ccg | 1658 |
| Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro | |
| 415 420 425 | |
| ggg atc ttc aag tcg acg tgc aag atc gac atc acg tgg ttc ccc ttc | 1706 |
| Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe | |
| 430 435 440 445 | |
| gat gac cag cgg tgc gag atg aag ttc ggc agt tgg acc tac gac gga | 1754 |
| Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly | |
| 450 455 460 | |

| | |
|---|------|
| ttc cag ctg gat tta caa tta caa gat gaa act ggc ggt gat atc agc | 1802 |
| Phe Gln Leu Asp Leu Gln Leu Gln Asp Glu Thr Gly Gly Asp Ile Ser | |
| 465 470 475 | |
| agt tac gtg ctc aac ggc gag tgg gaa cta ctg ggt gtg ccc ggc aaa | 1850 |
| Ser Tyr Val Leu Asn Gly Glu Trp Glu Leu Leu Gly Val Pro Gly Lys | |
| 480 485 490 | |
| cgt aac gag atc tat tac aac tgc tgc ccg gaa ccc tat ata gac atc | 1898 |
| Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile | |
| 495 500 505 | |
| acc ttc gcc atc atc atc cgc cga cga aca ctg tac tat ttc ttc aac | 1946 |
| Thr Phe Ala Ile Ile Ile Arg Arg Arg Thr Leu Tyr Tyr Phe Phe Asn | |
| 510 515 520 525 | |
| ctg atc ata cct tgt gta ctg att gcc tcc atg gcc ttg ctc gga ttc | 1994 |
| Leu Ile Ile Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe | |
| 530 535 540 | |
| acc ctg ccg cca gat tcg ggt gaa aaa tta tcg ctg ggt gtt acc atc | 2042 |
| Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile | |
| 545 550 555 | |
| ttg ctc tcg ctg acc gtg ttt ctg aat atg gtt gcc gag aca atg ccg | 2090 |
| Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro | |
| 560 565 570 | |
| gct act tcc gat gcg gtg cca ttg tgg ata cgc atc gtg ttt ttg tgc | 2138 |
| Ala Thr Ser Asp Ala Val Pro Leu Trp Ile Arg Ile Val Phe Leu Cys | |

575

580

585

tgg ctg cca tgg ata ttg cga atg agt cgc cca gga cga ccg ctg atc 2186

Trp Leu Pro Trp Ile Leu Arg Met Ser Arg Pro Gly Arg Pro Leu Ile

590

595

600

605

cta gag ttc ccg acc acg ccc tgt tgc gac aca tcc tcc gag cgg aag 2234

Leu Glu Phe Pro Thr Thr Pro Cys Ser Asp Thr Ser Ser Glu Arg Lys

610

615

620

cac cag ata ctc tcc gac gtt gag ctg aaa gag cgc tgc tgc aaa tgc 2282

His Gln Ile Leu Ser Asp Val Glu Leu Lys Glu Arg Ser Ser Lys Ser

625

630

635

ctg ctg gcc aac gta cta gac atc gat gat gac ttc cgg cac aat tgt 2330

Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Asn Cys

640

645

650

cgc ccc atg acg ccc ggc gga aca ctg cca cac aac ccg gct ttc tat 2378

Arg Pro Met Thr Pro Gly Gly Thr Leu Pro His Asn Pro Ala Phe Tyr

655

660

665

cgc acg gtt tat gga caa ggc gac gat ggc agc att ggg cca att ggc 2426

Arg Thr Val Tyr Gly Gln Gly Asp Asp Gly Ser Ile Gly Pro Ile Gly

670

675

680

685

agc acc cga atg ccg gat gcg gtc acc cat cat acg tgc atc aaa tca 2474

Ser Thr Arg Met Pro Asp Ala Val Thr His His Thr Cys Ile Lys Ser

690

695

700

tca act gaa tat gaa tta ggt tta atc tta aag gaa att cgc ttt ata 2522

Ser Thr Glu Tyr Glu Leu Gly Leu Ile Leu Lys Glu Ile Arg Phe Ile

705

710

715

act gat cag cta cgt aaa gat gac gag tgc aat gac att gcc aat gat 2570

Thr Asp Gln Leu Arg Lys Asp Asp Glu Cys Asn Asp Ile Ala Asn Asp

720

725

730

tgg aaa ttt gca gct atg gtc gtt gac aga ctg tgc ctt atc ata ttc 2618

Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys Leu Ile Ile Phe

735

740

745

aca atg ttc gca ata tta gcc aca ata gct gta cta cta tcg gca cca 2666

Thr Met Phe Ala Ile Leu Ala Thr Ile Ala Val Leu Leu Ser Ala Pro

750

755

760

765

cat att att gtc tcg tagccatatg ggcgagggtgg ttattgttat tggttttatt 2721

His Ile Ile Val Ser

770

ataaaatcaa tttgttaatt attaaattaa taacgaaact ctttaagtaa attaaaacta 2781

aaaagacact aaaaaagcac aaaaaaatag gaaaatacat gataaaaccc atgaactaaa 2841

taatacatcc aagaaaaacc aaaacaaaaa aaaaaaaaaa aaaaa

2886

<210> 3

<211> 3701

<212> DNA

<213> *Heliothis virescens*

<220>

<221> CDS

<222> (335)..(1822)

<400> 3

ggcacgagcc gctgccccac ggtcggccgc actccgctga acaacaatgc tcaaaaacac 60

gccgtgactc cacacacatc cctcggcgcc agtaggcgat gtttgaggat cggacggcac 120

gcgtggccgt cggcgagcgg tcgtgaacaa gttgcataca tatgaaaacc gtaaaaagat 180

tgaattttta gccgatcgtg ttcgatagat cctaatagag aagcggggagt gcggcgtttg 240

gtaggcgggg gtcgagtcgc gcggtcgggg gaaatggcgc ggcgcggggc ggcggcggcg 300

gcggcgcgcg gcgcggcggc gtcgcgcgcg tgac atg ggc ggg cgg gcg cgc cgc 355

Met Gly Gly Arg Ala Arg Arg

1

5

tcg cac ttg gcg gcg ccc gcg ggc ctg ctg ctg ctg tgc ctg ctc 403

Ser His Leu Ala Ala Pro Ala Gly Leu Leu Leu Leu Leu Cys Leu Leu

10

15

20

tgg ccg agg ggg gca cgc tgc ggg tac cac gag aag cgg cta ctg cac 451

Trp Pro Arg Gly Ala Arg Cys Gly Tyr His Glu Lys Arg Leu Leu His

25

30

35

cac cta ttg gac cac tac aac gta ctg gag agg ccc gtc gtc aac gag 499
 His Leu Leu Asp His Tyr Asn Val Leu Glu Arg Pro Val Val Asn Glu
 40 45 50 55

agc gac ccg ctg cag ctc tcc ttc ggc ctc acg ctc atg cag atc atc 547
 Ser Asp Pro Leu Gln Leu Ser Phe Gly Leu Thr Leu Met Gln Ile Ile
 60 65 70

gac gtg gac gag aag aac cag ctt tta ata aca aac atc tgg cta aaa 595
 Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr Asn Ile Trp Leu Lys
 75 80 85

cta gag tgg aat gat atg aac ttg agg tgg aac act tca gat ttc ggc 643
 Leu Glu Trp Asn Asp Met Asn Leu Arg Trp Asn Thr Ser Asp Phe Gly
 90 95 100

ggg gtc aaa gat tta aga gtg cca ccc cac aga cta tgg aaa cca gac 691
 Gly Val Lys Asp Leu Arg Val Pro Pro His Arg Leu Trp Lys Pro Asp
 105 110 115

gtc ctt atg tac aac agc gcg gac gaa ggg ttc gac agc acg tat cca 739
 Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe Asp Ser Thr Tyr Pro
 120 125 130 135

acg aac gtg gtg gtg cgg aac aac ggc tcg tgt ctg tac gtg ccg ccc 787
 Thr Asn Val Val Val Arg Asn Asn Gly Ser Cys Leu Tyr Val Pro Pro
 140 145 150

ggc atc ttc aag agc acc tgc aag atc gac atc acc tgg ttc ccc ttc 835
 Gly Ile Phe Lys Ser Thr Cys Lys Ile Asp Ile Thr Trp Phe Pro Phe

| | | | |
|---|-----|-----|------|
| 155 | 160 | 165 | |
| gac gac caa cga tgc gag atg aag ttt ggc agc tgg act tat gat ggt | | | 883 |
| Asp Asp Gln Arg Cys Glu Met Lys Phe Gly Ser Trp Thr Tyr Asp Gly | | | |
| 170 | 175 | 180 | |
| tat cag ttg gat cta caa cta cag gat gaa ggg ggc gga gat ata agc | | | 931 |
| Tyr Gln Leu Asp Leu Gln Leu Gln Asp Glu Gly Gly Gly Asp Ile Ser | | | |
| 185 | 190 | 195 | |
| agt ttt gtc acg aat ggc gaa tgg gag tta ata gga gtc ccc ggc aag | | | 979 |
| Ser Phe Val Thr Asn Gly Glu Trp Glu Leu Ile Gly Val Pro Gly Lys | | | |
| 200 | 205 | 210 | 215 |
| cgc aac gag atc tac tac aac tgt tgt ccg gag cca tac atc gac atc | | | 1027 |
| Arg Asn Glu Ile Tyr Tyr Asn Cys Cys Pro Glu Pro Tyr Ile Asp Ile | | | |
| 220 | 225 | 230 | |
| acg ttt gcg gtg gtg atc cgg agg aaa acg ctc tac tac ttc ttc aat | | | 1075 |
| Thr Phe Ala Val Val Ile Arg Arg Lys Thr Leu Tyr Tyr Phe Phe Asn | | | |
| 235 | 240 | 245 | |
| ctg atc gtg ccc tgc gtg ctc atc gcc tcc atg gct cta ttg ggg ttc | | | 1123 |
| Leu Ile Val Pro Cys Val Leu Ile Ala Ser Met Ala Leu Leu Gly Phe | | | |
| 250 | 255 | 260 | |
| acc ttg cct cca gac tcc gga gaa aag ttg tct tta ggt gtg acg ata | | | 1171 |
| Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Ser Leu Gly Val Thr Ile | | | |
| 265 | 270 | 275 | |
| tta ctg tcg ttg acg gtg ttc ctc aac atg gtg gcg gag acg atg cca | | | 1219 |

| | |
|---|-------------|
| Leu Leu Ser Leu Thr Val Phe Leu Asn Met Val Ala Glu Thr Met Pro | |
| 280 | 285 290 295 |
| gcg acg tcg gac gcc gtg ccc ttg ctc ggc acc tac ttc aac tgc atc | 1267 |
| Ala Thr Ser Asp Ala Val Pro Leu Leu Gly Thr Tyr Phe Asn Cys Ile | |
| 300 | 305 310 |
| atg ttc atg gtg gct tcc tcc gtc gtc tcc acc ata ctg atc ctc aac | 1315 |
| Met Phe Met Val Ala Ser Ser Val Val Ser Thr Ile Leu Ile Leu Asn | |
| 315 | 320 325 |
| tac cac cac cgg cac gca gac act cac gaa atg agt gat tgg att cgt | 1363 |
| Tyr His His Arg His Ala Asp Thr His Glu Met Ser Asp Trp Ile Arg | |
| 330 | 335 340 |
| tgc gtg ttc ctt tat tgg ctg ccg tgg gtg ctg cgc atg tca cgg ccc | 1411 |
| Cys Val Phe Leu Tyr Trp Leu Pro Trp Val Leu Arg Met Ser Arg Pro | |
| 345 | 350 355 |
| ggc tcg gcg acg acg ccg ccg ccg gcg cgc gta cct ccg ccg ccg gac | 1459 |
| Gly Ser Ala Thr Thr Pro Pro Pro Ala Arg Val Pro Pro Pro Pro Asp | |
| 360 | 365 370 375 |
| ctg gag ctg cgc gag cgc tcc tcc aag tcg ctc cta gcg aac gtg ctc | 1507 |
| Leu Glu Leu Arg Glu Arg Ser Ser Lys Ser Leu Leu Ala Asn Val Leu | |
| 380 | 385 390 |
| gac atc gat gac gac ttc cgc cac ccg caa gcg cag cag ccg caa tgc | 1555 |
| Asp Ile Asp Asp Asp Phe Arg His Pro Gln Ala Gln Gln Pro Gln Cys | |
| 395 | 400 405 |

tgc cga tac tac agg ggg ggt gag gag aat ggc gcg ggg ttg gcg gcg 1603
 Cys Arg Tyr Tyr Arg Gly Gly Glu Glu Asn Gly Ala Gly Leu Ala Ala
 410 415 420

cac agt tgc ttc ggt gtc gac tac gag ctc tcc ctc att ctg aag gag 1651
 His Ser Cys Phe Gly Val Asp Tyr Glu Leu Ser Leu Ile Leu Lys Glu
 425 430 435

att aga gtc atc aca gat cag atg cgc aag gac gac gaa gat gcg gac 1699
 Ile Arg Val Ile Thr Asp Gln Met Arg Lys Asp Asp Glu Asp Ala Asp
 440 445 450 455

att tcg cgc gac tgg aag ttc gcc gcc atg gtc gtg gac aga ctg tgc 1747
 Ile Ser Arg Asp Trp Lys Phe Ala Ala Met Val Val Asp Arg Leu Cys
 460 465 470

ctt att atc ttt acc ctg ttc aca atc atc gcc acg cta gcc gtg ctg 1795
 Leu Ile Ile Phe Thr Leu Phe Thr Ile Ile Ala Thr Leu Ala Val Leu
 475 480 485

ctg tcc gcg cca cac atc atg gtg tcg tagcgacccg cccgcttgcg 1842
 Leu Ser Ala Pro His Ile Met Val Ser
 490 495

gatacgcatg cgaaaagtgc tgtgataccg cgaatatattg ttaagttgtg atgagcgaag 1902

tggcgcggaac ggtgacgccg cggcgtcgga gttgccgccg cctgcctcgc cgccccgcgc 1962

cccctgtaga cataagttac cgctgactgc caaccctgta cgttcaacaa ataactgccc 2022

atccgactaa cgtcttttat ccccttgaaa aattcagcga ttgtgtaccc ctttcttcca 2082

agaatacaat gacaaatggt cgtcacgctc agtggaatca atcccgtact cttcgcccga 2142

tatttccctt agggatatgtc acgagtttga atgagcggtt ccgtatcaga cgttccgtcc 2202

ccggaacggt cgtcccctgc gataaagtgg cagtacgtgc tatacaggca cttaaggccg 2262

ccacgccacg gcgccgcggt gcgctcgggc cgcgaacccg cgaccctcac cgctgcaagt 2322

ggccaccac tagacaagac tgcggcagaa aatatttgca caaaaacgtc ttcctttctta 2382

ccgatgaacg acctgattcg catttaaaat taaactttgt tagaacttct tcgattcttg 2442

aatctattg tacagtttag agtttgggcg gtgaaacaat ggccctttgt ttccttcttg 2502

ttcgattcca tgaatcgtgg ttataatccc tagttttatt ttcggatata tttgtgtcag 2562

tagctagtat agaactttac aaacaatggt gattcaattg gtacagggtg tgatatgcct 2622

cgttgtgaac ggggccgata ttgttataaa tggtaaaata cccatggcta tagcttaata 2682

aatcgttcgt taaaagttgt agttaaaca atattatattt aataaagtca tatctgggtc 2742

ttccggaacg actttttaca ataattaaat tacatattaa taccacgttt gtactttctt 2802

ccatacagtt acagtaattc gtatgctgaa aataatatta gcttgtaaaa ttttcttctt 2862

cgaaaattta ttcaaacaga tgcgaccatc gtttcaaaca tttacatgta atatagaact 2922

cattttataa gatatacaac attttataag tacaagaagt tgtaacatga accggttttt 2982

--

cgttacatag aggggtataac acaaaggtgc ctacatattg acagatgcga agcacgatca 3042

gttgataagc acaggtacac tatatcctga catccgacag tccgtgccgt cgtctgccac 3102

actcggaac attcgacagt tcagtttact gtcgcccat catcgattgt taagtttgtt 3162

gttctaactc atcgattca ttccattcaa aaacattgta aacctctcaa ggggaaaacg 3222

tggtgtaaac agtgagagtg cgcgggtaca accgacacgc gaatgtaccc tcgcaaggct 3282

cctgtaatgt tttcctcttc cgagggtgtg ctgagagtaa tcttagacgg tccgatggaa 3342

gttgcggaac ggatatgatt acaagtcaat gtttttaagt catccgttta tttattgtta 3402

tatcttctta ccattcgcta gaggttgtgt gacgacccgg acggtgggcg ccgcaaccg 3462

cacacgcggg gttccatctt tgtattagat ggaagttgtg cggcatctct ccgtcggcaa 3522

tgggacaacc cgttgtcccc aacatttgtt caattgttag ggtaactct gaattgcact 3582

ttgtttatta aatataaacg aatgaaacaa aaaaaaaaaa aaaaaactcg agagtacttc 3642

tagagcggcc gcgggcccat cgattttcca cccgggtggg gtaccargta agtgtaccc 3701

<210> 5

<211> 3109

<212> DNA

<213> *Heliothis virescens*

<220>

<221> CDS

<222> (95)..(1597)

<400> 5

ggcacgagcc ggccgcacgt tgtcccaggc cgcattgagcg cgccggcgtg ctagcgcagc 60

gtgcgcgggt gtggtatgcc cgcgcgtcgc cgct atg gcc cct atg ttg gcg gcc 115

Met Ala Pro Met Leu Ala Ala

1

5

ttg gcg ctg ctg gct ttg ctg ccc gta tcg gag caa ggt cct cac gag 163

Leu Ala Leu Leu Ala Leu Leu Pro Val Ser Glu Gln Gly Pro His Glu

10

15

20

aag aga ctc ctg aac gcg ttg ctg gcg aac tac aac acc ctg gag cga 211

Lys Arg Leu Leu Asn Ala Leu Leu Ala Asn Tyr Asn Thr Leu Glu Arg

25

30

35

ccg gtg gcc aac gag agc gaa ccg cta gag gtc agg ttc ggc ttg acc 259

Pro Val Ala Asn Glu Ser Glu Pro Leu Glu Val Arg Phe Gly Leu Thr

40

45

50

55

ttg cag caa atc att gac gtg gac gag aag aat caa cta ctt ata acc 307

Leu Gln Gln Ile Ile Asp Val Asp Glu Lys Asn Gln Leu Leu Ile Thr

60

65

70

aat ata tgg ctg tct ttg gag tgg aat gac tac aac ctg agg tgg aac 355

Asn Ile Trp Leu Ser Leu Glu Trp Asn Asp Tyr Asn Leu Arg Trp Asn

75

80

85

gac agc gag tat ggc ggg gtc aag gac ctc agg atc acg ccc aac aag 403

Asp Ser Glu Tyr Gly Gly Val Lys Asp Leu Arg Ile Thr Pro Asn Lys

90

95

100

ttg tgg aag ccg gac gtc ctt atg tat aat agt gct gac gag ggt ttt 451

Leu Trp Lys Pro Asp Val Leu Met Tyr Asn Ser Ala Asp Glu Gly Phe

105

110

115

gac ggg acc tac cag acc aac gtg gtg gtc aga agc ggc ggc agt tgc 499

Asp Gly Thr Tyr Gln Thr Asn Val Val Val Arg Ser Gly Gly Ser Cys

120

125

130

135

ctg tac gtg cca cct ggc ata ttc aag agc aca tgc aag atg gac atc 547

Leu Tyr Val Pro Pro Gly Ile Phe Lys Ser Thr Cys Lys Met Asp Ile

140

145

150

gcg tgg ttt ccc ttc gac gac caa cac tgt gat atg aag ttc ggt agc 595

Ala Trp Phe Pro Phe Asp Asp Gln His Cys Asp Met Lys Phe Gly Ser

155

160

165

tgg aca tat gac ggc aat cag ttg gat ctg gtg cta aaa gat gag gca 643

Trp Thr Tyr Asp Gly Asn Gln Leu Asp Leu Val Leu Lys Asp Glu Ala

170

175

180

| | |
|---|------|
| ggc ggc gat cta tcg gac ttc ata aca aat ggg gag tgg tat cta ata | 691 |
| Gly Gly Asp Leu Ser Asp Phe Ile Thr Asn Gly Glu Trp Tyr Leu Ile | |
| 185 | 190 |
| 195 | |
| gga atg cca ggc aaa aag aac aca ata aca tac gcg tgc tgc ccc gag | 739 |
| Gly Met Pro Gly Lys Lys Asn Thr Ile Thr Tyr Ala Cys Cys Pro Glu | |
| 200 | 205 |
| 210 | 215 |
| ccc tac gtg gac gtc acc ttc acc atc atg ata aga aga cga acc ttg | 787 |
| Pro Tyr Val Asp Val Thr Phe Thr Ile Met Ile Arg Arg Arg Thr Leu | |
| 220 | 225 |
| 230 | |
| tac tac ttc ttc aac ctg atc gtc ccg tgc gtg ctg atc tca tcg atg | 835 |
| Tyr Tyr Phe Phe Asn Leu Ile Val Pro Cys Val Leu Ile Ser Ser Met | |
| 235 | 240 |
| 245 | |
| gca ctc ctc ggc ttc aca ctg cca cca gac tcc gga gag aaa ctc aca | 883 |
| Ala Leu Leu Gly Phe Thr Leu Pro Pro Asp Ser Gly Glu Lys Leu Thr | |
| 250 | 255 |
| 260 | |
| ctt gga gtc act att ctt cta tcg ctg acg gtg ttc ctc aac ctg gta | 931 |
| Leu Gly Val Thr Ile Leu Leu Ser Leu Thr Val Phe Leu Asn Leu Val | |
| 265 | 270 |
| 275 | |
| gcc gag acc ctg cca cag gtc tcc gac gct atc ccc ctg tta ggg acg | 979 |
| Ala Glu Thr Leu Pro Gln Val Ser Asp Ala Ile Pro Leu Leu Gly Thr | |
| 280 | 285 |
| 290 | 295 |
| tac ttc aat tgc atc atg ttc atg gta gcg tcg tct gtg gta ctg act | 1027 |
| Tyr Phe Asn Cys Ile Met Phe Met Val Ala Ser Ser Val Val Leu Thr | |
| 300 | 305 |
| 310 | |

gtg gtg gta ctc aat tac cac cat cga aca gct gat ata cat gaa atg 1075
 Val Val Val Leu Asn Tyr His His Arg Thr Ala Asp Ile His Glu Met
 315 320 325

cca cag tgg ata aaa tca gta ttc cta caa tgg ttg cca tgg ata ctg 1123
 Pro Gln Trp Ile Lys Ser Val Phe Leu Gln Trp Leu Pro Trp Ile Leu
 330 335 340

cga atg tcg agg cca ggg aag aag atc acc agg aag act ata atg atg 1171
 Arg Met Ser Arg Pro Gly Lys Lys Ile Thr Arg Lys Thr Ile Met Met
 345 350 355

aac acg agg atg agg gag ctg gaa ctg aag gag agg tcg tcg aag tcc 1219
 Asn Thr Arg Met Arg Glu Leu Glu Leu Lys Glu Arg Ser Ser Lys Ser
 360 365 370 375

ttg ctg gcg aat gtt cta gat att gat gat gac ttc aga cac ggc cct 1267
 Leu Leu Ala Asn Val Leu Asp Ile Asp Asp Asp Phe Arg His Gly Pro
 380 385 390

ccg cct cct aac agt act gcc tcg acc ggg aat ttg gga cct ggg tgc 1315
 Pro Pro Pro Asn Ser Thr Ala Ser Thr Gly Asn Leu Gly Pro Gly Cys
 395 400 405

tca ata ttc cgc acg gat ttc cgt cgg tcg ttc gtc cgt ccg tcc acg 1363
 Ser Ile Phe Arg Thr Asp Phe Arg Arg Ser Phe Val Arg Pro Ser Thr
 410 415 420

atg gaa gac gtg ggc ggc ggg ctg ggt agc cac cat cgc gag ctg cac 1411
 Met Glu Asp Val Gly Gly Gly Leu Gly Ser His His Arg Glu Leu His

425

430

435

ctc ata ctg aga gag ctg cag ttc atc acg gcc agg atg aag aag gct 1459

Leu Ile Leu Arg Glu Leu Gln Phe Ile Thr Ala Arg Met Lys Lys Ala

440

445

450

455

gat gag gaa gcc gag ctg atc agc gac tgg aag ttt gct gcg atg gtt 1507

Asp Glu Glu Ala Glu Leu Ile Ser Asp Trp Lys Phe Ala Ala Met Val

460

465

470

gtt gat agg ttt tgc ctg ttc gtg ttc aca ctt ttc aca atc atc gcg 1555

Val Asp Arg Phe Cys Leu Phe Val Phe Thr Leu Phe Thr Ile Ile Ala

475

480

485

aca gta gct gtc ctg tta tcg gca ccg cat atc atc gtg caa 1597

Thr Val Ala Val Leu Leu Ser Ala Pro His Ile Ile Val Gln

490

495

500

tgaaccaacc actgagccgg caactccggc gcatgaatga gagaaataat tattagatcg 1657

ccgatttgta attataattg ataatgtaat taaattaaat acgtgggttga aacgcacacg 1717

tctccataac aaagtcttaa gacattaaat tatgataaat ttacatattg tagttaagtc 1777

gagtgttgat ggaaatttta gccggcgcaa ggagtttcgt gaaggtctgt atatattttt 1837

tcttattggt gtatattgta tcgttggttca tgttttcttt caggaagtga gctttgtact 1897

gtttgtttct tcgatggcag gtgcacttca gttcaggctg aaatttccat taacatttat 1957

ttaaacaat gtgatgttga ctaggatgtt atacagataa atgttgacgt gtataatttg 2017

ttaaaataaa caatattaat tactattact aaacgatatt ataaacgaag tactaacgag 2077

ggttacttta atggaagaa cgctaagctg gcacagagtt gcattaattt gaaaaaagaa 2137

attacggaaa aaagtttatt gaaaattgaa ctttttgga ggaaagtaac gtttgatcaa 2197

aaaagtttgt aaaacgaaag ttcggttctg cgccaatact ggaattaaaa ttctcgtaaa 2257

tattagggaa aagaaggtcc tttaaaca aaagatttgaa ccggcatcct tttacaagt 2317

aatgagggat cacagatgat gacaaaaaac cttagggat ataagtaatg tacataatgg 2377

atcaaatac ggtagagtca agaatagtta acgatttaag attattccat tcgatattaa 2437

aattcgatta gcgattgtcg ctgcgtctac ttgatacat atcgatttga atcgatattg 2497

tataaattta gatagatcgg acattagtaa tgagtatgga cgttttaatt tttaaaaaag 2557

aatgtactac gaagattaaa tccaggaatt gttaaacagt tatggaattg ataagaaatc 2617

aacaattaat acggaaccaa aggtagacta ggtgtagcat caggagattg aattaaaaca 2677

taaattagga ccgacttaaa tggaacttgc gagtgtattg ataacttttt aatttaaaaa 2737

ctcattgtcg attaaatgga gaataacttt tgatctctcg tatcgataaa tgctcactta 2797

actatcgata gcgtaatat ataatgtta gtatatcgat atgggagtaa gtcactagca 2857

tcagaaatag tcattaatta ggaatcggtt tgtgttaatg ttatgcttag cgaaaatatt 2917

acaatgctgt tgatattcact aaccatcacg taaccatatt gataaaatgt aaatacagaa 2977

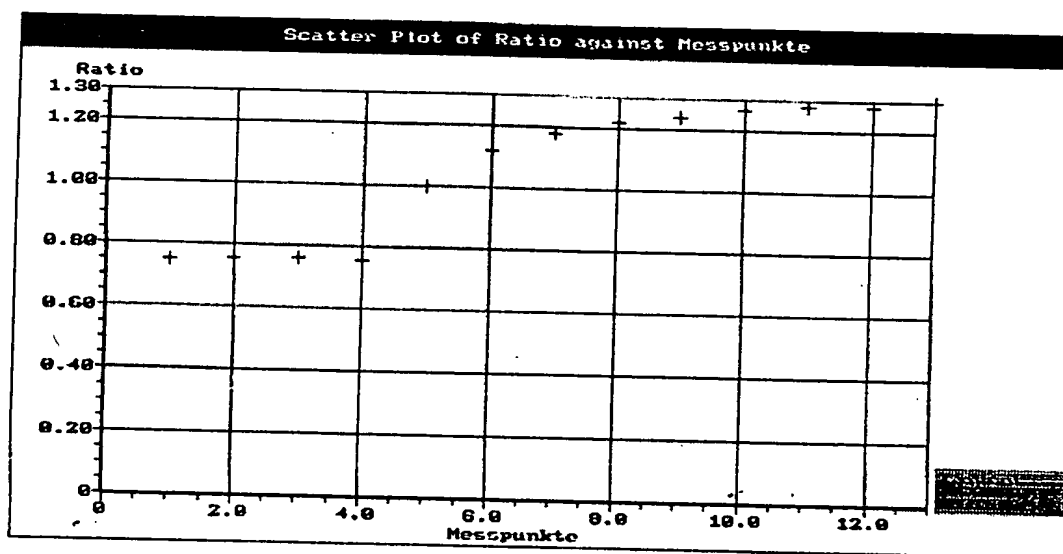
tattgcggtg tgtatttgta tataaatttt agaaaaaaaa aaaaaaaaaa aactcgagag 3037

tacttctaga gcggccgcgg gcccatcgat tttccaccg ggtggggtac caggtaagtg 3097

taccaattc gc

3109

Figure 1



Measuring Points

MESS PUNKTE

↑ ↑
MEASURING - POINT

DATA - POINT